

Investigations in Instructed Second Language Acquisition



Studies on Language Acquisition  
25

*Editor*

Peter Jordens

Mouton de Gruyter  
Berlin · New York

# Investigations in Instructed Second Language Acquisition

*edited by*

Alex Housen

Michel Pierrard

Mouton de Gruyter  
Berlin · New York

Mouton de Gruyter (formerly Mouton, The Hague)  
is a Division of Walter de Gruyter GmbH & Co. KG, Berlin.

⊗ Printed on acid-free paper which falls within the guidelines  
of the ANSI to ensure permanence and durability.

*Library of Congress Cataloging-in-Publication Data*

Investigations in instructed second language acquisition / edited  
by Alex Housen, Michel Pierrard.

p. cm. — (Studies on language acquisition ; 25)

Includes bibliographical references and index.

ISBN 3-11-017970-9 (cloth : alk. paper)

1. Second language acquisition. 2. Language and languages — Study and teaching. I. Housen, Alex, 1964—  
II. Pierrard, M. III. Series.

P118.2.I583 2005

418'.0071—dc22

2005003499

ISBN 3 11 017970 9

*Bibliographic information published by Die Deutsche Bibliothek*

Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie;  
detailed bibliographic data is available in the Internet at <<http://dnb.ddb.de>>.

© Copyright 2005 by Walter de Gruyter GmbH & Co. KG, D-10785 Berlin.

All rights reserved, including those of translation into foreign languages. No part of this book  
may be reproduced in any form or by any means, electronic or mechanical, including photocopy,  
recording, or any information storage and retrieval system, without permission in writing from  
the publisher.

Cover design: Sigurd Wendland, Berlin.

Printed in Germany.

## Preface

This volume grew out of a Colloquium on “*Instructed Second Language Learning/L’appropriation d’une langue seconde en milieu guidé*” organized by the editors of this volume in Brussels in August 2004. The purpose of the colloquium was to assess the impact of the ‘new wave’ of multidisciplinary research on second language acquisition (SLA) in instructional settings for theory, methodology, and practice. This volume captures the spirit of the conference participants’ free exchange of ideas and goes substantially beyond that event. We asked the presenters to update, revise and expand their papers, and we also invited additional contributions, in an effort to present complementary, multiple perspectives on the investigation of instructed second language acquisition (ISLA). All the papers submitted for inclusion in this book went through a rigorous review process. As it stands, the volume brings together recent work, both empirical and theoretical, on aspects of learning, processing and use of a second language in a variety of instructional settings, including foreign language classes, intensive second language classes, immersion programmes, and content-and-language-integrated learning environments. The introductory chapter provides a brief overview of the assumptions that underlie research on ISLA, followed by a discussion of factors that may determine the process of instructed second language acquisition and its possible outcomes. The next seventeen chapters report theoretical and empirical work in this field. Data-based studies in this book deal with the acquisition of specific linguistic phenomena (e.g. verb and noun morphology, lexicon, clause structures) in a range of target languages (e.g. English, French, German, Russian). Several of the chapters deal with the role of form-focused and meaning-focused instruction in L2 learning, but other issues such as the role of cross-linguistic influence, awareness, implicit and explicit processing mechanisms, memory and the properties of classroom input are also discussed. Although the interest in instruction in this volume is acquisitional rather than pedagogical, and all the chapters address theoretical questions, several also consider pedagogical implications for language educators. As such we believe that this volume will prove a valuable resource for researchers in SLA, psycholinguistics, linguistics and language pedagogy.

Alex Housen

Michel Pierrard

## **Acknowledgements**

We have been very lucky to obtain contributions from a wide range of distinguished scholars in major centres of research in both Europe and North America. For reasons beyond our control, this volume appears later than we had intended. We are grateful to all contributors for their patience and for their efforts given to the individual chapters in this book. We also gratefully acknowledge the thorough and constructive work of twenty-three anonymous reviewers. We have also greatly appreciated the support and expert guidance of the series editor, Peter Jordens, and of Ursula Kleinhenz at Mouton de Gruyter. Finally, we especially wish to thank Frank Benno Junghanns for his expertise in copy editing.

## Contents

Preface	v
Investigating Instructed Second Language Acquisition <i>Alex Housen and Michel Pierrard</i>	1
<b>Investigating cognitive and processing mechanisms in instructed SLA</b>	
Instructed learners' fluency and implicit/explicit language processes <i>Liz Temple</i>	31
Psycholinguistic aspects of gender acquisition in instructed GFL learning <i>Barbara Menzel</i>	51
Language analytic ability and oral production in a second language: Is there a connection? <i>Leila Ranta</i>	99
Formal instruction and the acquisition of verbal morphology <i>Kira Gor and Tatiana Chernigovskaya</i>	131
<b>Investigating the role and effects of form-focused instruction</b>	
Teaching marked linguistic structures – more about the acquisition of relative clauses by Arab learners of English <i>Ahlem Ammar and Patsy M. Lightbown</i>	167
The importance of form/meaning mappings in explicit form-focussed instruction <i>Nina Spada, Patsy M. Lightbown and Joanna L. White</i>	199

Rule complexity and the efficacy of explicit grammar instruction <i>Alex Housen, Michel Pierrard and Siska Van Daele</i>	235
Focus on forms as a means of improving accurate oral production <i>Ronald Sheen</i>	271
Instructed Second Language Vocabulary Learning: The fault in the ‘default hypothesis’ <i>Batia Laufer</i>	311
<b>Investigating the role and effects interaction and Communication-Focused Instruction</b>	
Negative feedback and learner uptake in analytic foreign language teaching <i>Katja Lochtmann</i>	333
Noticing and the role of interaction in promoting language learning <i>Folkert Kuiken and Ineke Vedder</i>	353
Interactional strategies for interlanguage communication: Do they provide evidence for attention to form? <i>María del Pilar García Mayo</i>	383
Assessment of the role of communication tasks in the development of second language oral production skills <i>Peter Griggs</i>	407
Language learning in content-based instruction <i>Heini-Marja Järvinen</i>	433
Effects of teacher discourse on learner discourse in a second language classroom <i>Tsuyoshi Kida</i>	457



**Comparing the effects of instructed and naturalistic  
L2 acquisition contexts**

Second language acquisition in a study abroad context: A comparative investigation of the effects of study abroad and foreign language instruction on the L2 learner's grammatical development <i>Martin Howard</i>	495
The effect of type of acquisition context on perception and self-reported use of swearwords in the L2, L3, L4 and L5 <i>Jean-Marc Dewaele</i>	531
Index	561
List of Contributors	567



# Investigating Instructed Second Language Acquisition

*Alex Housen and Michel Pierrard*

## 1. Introduction

Most SLA research makes a basic distinction between *uninstructed* (*naturalistic, spontaneous, unguided, untutored, informal*) second language acquisition (SLA) and *instructed* (or *guided, tutored, formal*) second language acquisition, according to whether the second language (L2) is learned through spontaneous communication in authentic social situations or under pedagogical guidance (Ellis 1985, 1994; Klein 1986; McLaughlin 1987; Larsen-Freeman & Long 1991)<sup>1</sup>. The practical validity of this distinction may seem unquestionable but it is unclear whether instructed and uninstructed SLA are really different processes and opinions in the literature on this point differ widely.

At one extreme, there is the view that instructed and uninstructed SLA are fundamentally different processes. The best-known proponent of this view is Krashen who proposed the dichotomy between uninstructed L2 *acquisition* versus instructed L2 *learning* and rejected the possibility of an interface between the two types (Krashen 1981, 1985). Instruction, Krashen argues, leads to learned conscious knowledge, which is available to the learner as a monitor for checking the form of utterances once they have been generated by the acquired subconscious knowledge system. Krashen rules out the possibility of instruction intervening in the acquisition process as it proceeds along some fixed natural order. Often the implication here is that the only ‘real’ SLA – or at least the only kind of SLA that merits the researcher’s attention – is uninstructed SLA; that is, SLA uncontaminated by the intervention and control exerted by the classroom, instructor or textbook. This bias is made explicit in the following quote: “What happens in school has very little to do with language learning. Language can’t be taught. It can only be learned. People learn language in spite of what goes on in the classroom” (cited in Wong-Fillmore 1989: 315).

At the other extreme, there are those who claim that SLA always involves the same basic processes, regardless of context: “It is difficult to imagine a situation in which the fundamental processes involved in learning a non-primary language would depend on the context in which the language is learned ... All learners have the capability of taking information from the input and organising it within the framework of their current linguistic system and modifying and restructuring that system” (Gass 1989: 498; see Felix 1981 or Bardovi-Harlig 2000, for similar views).

These two seemingly opposing views both in fact regard SLA as an essentially self-contained process that follows its own course, a process neither dependent on nor influenced by external factors. Thus, as Van Patten observes, “it is what [a] learner does that is common to all contexts which forms the core of SLA theory” (Van Patten 1990: 25).

Although extremes of opinion do exist, it would be true to say that most SLA researchers nowadays, including the contributors to the present volume, would consider them misguided or at least premature. Instead, SLA is typically considered to be a process which is open to the influence of instruction. What is not fully understood is exactly how, and to what extent, the process can be influenced. It was these crucial research questions which provided the starting point for the studies in this volume.

## 2. Investigating Instructed Second Language Acquisition

In order to find out if and how instructed and uninstructed SLA differ, we have to be more specific about what we understand by *instruction*. As will be apparent from the contributions to this volume, instruction is not a unitary concept and the term is used to mean different things depending on the theoretical perspective and research focus. For current purposes therefore, we define instruction as *any systematic attempt to enable or facilitate language learning by manipulating the mechanisms of learning and/or the conditions under which these occur*. This broad definition allows for a wide range of instructional approaches, methods, strategies, techniques, practices and activities, all of which can be applied in a wide range of settings (typically a classroom). Thus institutionalized forms of L2 instruction and methods of training are obviously included; but so are individualized L2 instruction, self-study, computer-assisted instruction and the use of audio-visual and electronic learning materials.

The study of Instructed Second Language Acquisition (henceforth *ISLA*) is motivated by several concerns.

First, ISLA merits our attention because it is an important social phenomenon. An ever-increasing number of people, particularly in the developed world, are learning a second language at least partially through instruction, mainly in the controlled environment of a classroom. Indeed, *instructed* L2 acquisition may well be the predominant mode of SLA, more so than *naturalistic* L2 acquisition. Consequently the study of ISLA has great descriptive value and ecological validity.

Secondly, the study of ISLA also has applied value, especially for language education. L2 learning and L2 teaching are both highly complex tasks that require much time, effort and resources from the learner, the instructor, and the community. Insights from ISLA research can reveal the complexity of these tasks and contribute to improvements in instructional practice.

Finally, the study of ISLA has theoretical value. It calls for a consideration of a wide range of theoretical issues pertaining to the nature of language, language learning, language knowledge and language processing, and the relationships between them. The SLA community needs answers to a range of fundamental questions, including the following:

1. What is the nature of the *learning mechanisms* involved in instructed SLA and how do they differ from the learning mechanisms at work in uninstructed SLA? For instance, is there a basis (cognitive, linguistic or neurological) for distinctions commonly made between *acquisition* and *learning*, between *implicit* and *explicit learning*?
2. What is the nature of the *L2 knowledge* that instructed L2 learners develop, and how does it differ from the L2 knowledge that develops in naturalistic, uninstructed SLA? This question takes in issues such as the nature of metalinguistic knowledge and the status of distinctions such as implicit versus explicit language knowledge and procedural versus declarative knowledge.
3. What is the nature of *L2 performance and L2 processing* in instructed L2 contexts? This question addresses the distinction between controlled versus automatic processing, the nature and role of monitoring and planning, and the effect of task type on L2 performance.
4. What is the nature of *L2 instruction*? Various types of instruction are currently distinguished in the literature, including consciousness raising, input flooding, input enhancement, focus-on-form, focus-on-

formS, instruction as the provision of comprehensible input or of positive and negative evidence, instruction as providing strategies for input processing, and so forth. What do these different forms of instruction have in common and how do they differ?

These issues are clearly related and cannot be investigated in isolation. We need to understand how they interact, or in other words:

5. What is the *relationship between instruction, acquisition, knowledge and processing in an L2 and how do they interact?* What variables intervene to constrain this interaction (e.g. age, cognitive maturity, motivation, first language background, the nature of the L2 features targeted by instruction)? In dealing with this more complex question, we need to address issues such as “What type of instruction leads to what kind of knowledge and how is this knowledge activated in subsequent L2 processing”?

Any attempt to find answers to these questions makes the lamentable inadequacy of a uni-disciplinary approach quickly apparent. Each major issue raises a whole host of related questions, as we will see below, where we examine in more detail just one of the issues raised above and one which many consider to be the core issue of ISLA research: the role of instruction in SLA.

### **3. Investigating the role of instruction**

The role and effects of instruction in SLA have always been controversial: Does instruction really enable SLA, or at least facilitate it? Language teachers have (perhaps obviously) always believed that instruction enables, or at least facilitates language acquisition, but SLA researchers have sometimes been less certain. Early reviews of research on the role of instruction on SLA found it useful to consider its effect in terms of the route, rate and end-state of L2 acquisition (e.g. Long 1983; Chaudron 1988; Ellis 1984, 1985, 1990; Harley 1988). The mainly descriptive research suggested that instruction could positively affect the rate and end-state of acquisition but not its route. In other words, instructed learners would progress faster and ultimately attain higher levels of proficiency than uninstructed L2 learners but both instructed and non-instructed learners would proceed through the same stages and sequences of acquisition, suggesting that instructional intervention is incapable of overriding certain ‘natural’ mechanisms and universal predispositions operative in SLA.

The findings of more experimental research offered a slightly different picture. For instance, early experimental research on the Teachability Hypothesis (Pienemann 1984, 1987a,b) suggested that the natural developmental sequence cannot be altered by instruction for elements of language whose acquisition is governed by universal processing constraints but that other, variational features of language can in principle be successfully taught at any stage of development. Similarly, research inspired by markedness theories suggested that instruction can lead learners to skip stages in a developmental sequence if the instruction is targeted at the more marked forms in an implicationally related hierarchy of forms such as relative clause structures (Gass 1982; Eckman, Bell & Nelson 1988; Ammar & Lightbown, this volume). More recent research on the nature and role of instruction, as reviewed in Norris & Ortega (2000) and Ellis (2001, 2002), paints an even more complex picture. This recent research suggests that any attempt to understand the role of instruction should not treat it as a unitary concept but rather as a cover for a wide range of activities and practices differing along a number of dimensions and potentially affecting different aspects of L2 learning, L2 competence and L2 performance.

In the following sections, we propose a framework for describing the role of instruction in SLA. This framework includes both (a) the nature of the effects of instruction on SLA and (b) the factors which mediate these effects and hence, the effectiveness of instruction.

### 3.1. Effects of instruction

A proper understanding of the role of instruction in SLA requires a certain clarity about the variegated effects which instruction may have on SLA. These effects can be envisaged in terms of (1) the basic dimensions of SLA, (2) the basic components of SLA, (3) the major processes of SLA and (4) the different types of knowledge which instructed L2 learners develop. These four factors are briefly elucidated below.

First, instruction can, at least in principle, affect any one of the three *basic dimensions* of the language learning process (Klein 1986; Ellis 1994):

- it may affect the *route of acquisition* (i.e. instructed learners may internalise the various features of the target language in a different order from non-instructed learners);

- it may affect the *rate* of language learning (i.e. accelerate or slow it down);
- it may affect *ultimate levels of attainment* and the ‘*end-state*’ of learning (i.e. instructed language learners may ultimately reach either higher or lower stages of development, or attain higher or lower levels of proficiency than non-instructed learners).

Secondly, in terms of the *basic components* of SLA (Klein 1986; Ellis 1994), instruction can be viewed as doing one or several of the following:

- instruction can provide learners with *exposure* to the target language (i.e. input and output opportunities) which is otherwise insufficiently available;
- instruction can influence learners’ *propensity* to use and learn the target language (e.g. by stimulating their motivation);
- instruction can trigger *learning processes and mechanisms* which are otherwise insufficiently activated (e.g. automatization processes, restructuring of linguistic representations).

Thirdly, for the purpose of describing the role of instruction, the third component listed above, L2 learning processes, can be envisaged as comprising three broad *types of processes*: *knowledge internalisation*, *knowledge modification* and *knowledge consolidation*. The goals and effects of instruction can be accordingly characterized as follows:

- instruction may enable learners to *internalize* new L2 knowledge (so that they become *more elaborate* L2 users with, for example, a richer vocabulary and more complex grammar);
- instruction may enable learners to *modify (restructure)* their L2 knowledge and performance, particularly the deviant, non-targetlike aspects of their knowledge and performance (so that they become *more accurate*);
- instruction may enable learners to *consolidate* their L2 knowledge (so that they can use the L2 with greater ease and for a wider range of tasks and functions, in short, so that they become *more fluent* language users).

An important issue here is the role of *consciousness* and *attention* in the process(es) of language learning. Recent research has defined consciousness as *awareness* and has argued that the acquisition of language knowledge involves the allocation of attentional resources to language features in the



input. Simply stated, “people learn about the things that they attend to and do not learn much about the things that they do not attend to” (Schmidt 2001: 30). Depending on the type and amount of attentional resources allocated, different levels of awareness are distinguished, ranging from *perception* and *detection*, to *noticing* and finally to *understanding* language features (Schmidt 1995; Robinson 1996; De Graaff 1997).<sup>2</sup> The distinction between these different mental operations is scalar, rather than categorical. The critical level of awareness for language learning is *noticing*: “intake is that part of the input that the learner notices” (Schmidt 1990: 139). Consciousness as awareness at the level of rule understanding is considered merely facilitative of attempts to learn. In this view then, the primary role of instruction is as a means for promoting noticing of relevant language features.

The noticing/understanding distinction is related to the distinction between *implicit* and *explicit* learning familiar from research in experimental psychology (e.g. Reber 1993). In contrast to Krashen who considered (implicit) *acquisition* and (explicit) *learning* as fundamentally different, other researchers view implicit and explicit learning as fundamentally similar processes, as both involve the allocation of attentional resources to input and both result in memorial representations of input features (Robinson 1996). The distinction between implicit and explicit learning is defined at the level of their different resultant knowledge bases, as determined by the conditions under which the learning occurs and the type of input provided.

This leads us to the fourth and final way in which the goals and effects of instruction can be envisaged, namely in terms of the types of language knowledge which it promotes. The most common distinctions in SLA research are between *implicit* and *explicit* knowledge and *declarative* and *procedural* knowledge. A survey of the literature on these two distinctions reveals considerable definitional discrepancies (cf. Ellis 1997; Johnson 1996). Implicit knowledge is often characterized as largely *intuitive* and abstract knowledge of language which is *acquired* subconsciously and *incidentally* (typically as a ‘by-product’ of engaging in authentic communication) and which is generally considered as the basis of unplanned, communicative language use. In contrast, explicit knowledge, broadly defined as knowledge *about* language, is a more conscious type of knowledge that is *learned intentionally*. Explicit knowledge can be broken down further into *analysed knowledge* and *metalinguistic knowledge*. Metalinguistic knowledge is verbalized knowledge about the structure and knowledge of language and of the theoretical constructs and technical or semi-technical terminology used to describe it. It is learned through deliberate and conscious (and often

conscientious) study and involves a higher form of awareness than analysed knowledge. Analysed knowledge refers to the extent to which learners are able to form a propositional mental representation of language features and rules. According to Bialystok (1994), analyzed knowledge is derived from implicit knowledge as learners begin to decode their implicit knowledge linguistically so that it becomes represented in a more analytic and symbolic form. Analysed knowledge typically manifests itself in problem-solving language tasks which require learners to pay focal attention to the choice of linguistic forms (as in a cloze task or grammaticality judgement task) though it can also manifest itself intermittently in naturally occurring language behaviour (e.g. in tasks involving decontextualised language use and the manipulation of complex formal schemata). Because this type of knowledge cannot be accessed rapidly, it is normally only activated when there is opportunity for reflection and language planning (*monitoring*).

As said earlier, other researchers distinguish between *declarative* versus *procedural* L2 knowledge, where the former is usually characterized as ‘knowing that’ and the latter as ‘knowing how’. Language learners may first represent a particular language feature as declarative knowledge in memory (e.g. in the form of a set of semantic networks) and then go on to *proceduralize* this declarative knowledge by converting it into mental procedures which can be rapidly activated for processing utterances in spontaneous language use. *Proceduralization* in this sense is similar to the process of *automatisation* and may be seen as part of the sub-process of *knowledge consolidation* mentioned earlier. Although the declarative-procedural distinction is sometimes equated with the implicit-explicit distinction, they are not necessarily the same. Ellis (2001), for example, allows for the possibility of both implicit and explicit knowledge to be available in declarative as well as proceduralized form.

It follows from the above that when one wants to describe the effects of instruction, one has to allow for the possibility that instruction can promote different types of L2 knowledge, including declarative implicit knowledge, procedural implicit knowledge, declarative explicit knowledge, procedural explicit knowledge, analytic explicit knowledge and metalinguistic knowledge.

### 3.2. Mediating factors

Whatever the nature of the effects, and the effectiveness of instruction for SLA, it seems reasonable to assume that they will be mediated by at least three factors, relating to the *how*, the *what* and the *who* of instruction: (a) the type of instruction provided, (b) the type of language features targeted for instruction, and (c) the type of learner at whom the instruction is targeted (De Graaff 1997; Norris & Ortega 2000; Ellis 2002).

Starting with the learner factor, it is generally assumed – though insufficiently demonstrated – that instruction will have different effects, and hence be more or less effective, depending on the individual learner’s age, cognitive maturity, cognitive style, motivation, personality, language learning aptitude and level of L2 proficiency at the time of instruction (cf. Skehan 1989, 2002; Ranta 2002; Sawyer & Ranta 2001; Larsen-Freeman & Long 1991).

In contrast to the impact of learner variables on the effectiveness of instruction, the mediating role of *type of instruction* has been the focus of much recent SLA research. An initial broad distinction can be made between *Communication-Focused Instruction* and *Form-Focused Instruction* (cf. Ellis 1999). Communication-Focused Instruction (CFI) aims to engage the learner in the active negotiation of meaning and the communicative exchange of authentic messages (Ellis 1999). The theoretical underpinnings of CFI derive from the Comprehensible Input Hypothesis (Krashen 1985) and the Comprehensible Output Hypothesis (Swain 1985). CFI assumes that language is best learned through the comprehension of input and through the noticing of form-function mappings which results from the learner’s own attempts to actively negotiate meaning in interaction (Ellis 1999). Form-Focused Instruction (FFI) refers to “any pedagogical effort used to draw the learner’s attention to language form [...]” (Spada 1997: 73). FFI is based on the assumption that certain features of language – grammatical structures but also lexical items, phonological and even sociolinguistic and pragmatic features – can go unnoticed in the input unless the learner’s attention is somehow drawn to them so that he reaches the critical level of awareness (*noticing*) for the features to be internalized (Sharwood Smith 1993; Schmidt 1995). FFI can take many forms. One way of classifying them is in terms of their degree of explicitness: from implicit instructional techniques such as input flooding, input enhancement techniques and recasts to increasingly more explicit techniques like controlled focused exercises, overt error correction and the presentation and discussion of metalinguistic rules (Sharwood Smith 1993). Table 1 contrasts a number of

the attributes associated with implicit and explicit forms of FFI (cf. Norris & Ortega 2000; DeKeyser 1995; Ellis 2001, 2002).

Table 1. Implicit and explicit forms of Form-Focused Instruction

Implicit FFI	Explicit FFI
<ul style="list-style-type: none"> <li>• <i>attracts</i> attention to target form</li> <li>• is delivered <i>spontaneously</i> (e.g. in an otherwise communication-oriented activity)</li> <li>• is unobtrusive (minimal interruption of communication of meaning)</li> <li>• presents target forms in context</li> <li>• makes no use of metalanguage</li> <li>• encourages free use of target form</li> </ul>	<ul style="list-style-type: none"> <li>• <i>directs</i> attention to target form</li> <li>• is <i>predetermined</i> and <i>planned</i> (e.g. as the main focus and goal of a teaching activity)</li> <li>• is obtrusive (interruption of communication of meaning)</li> <li>• presents target forms in isolation</li> <li>• uses metalinguistic terminology (e.g. rule explanation)</li> <li>• involves controlled practice of target form</li> </ul>

The distinction between implicit and explicit FFI covers the well-known distinction between respectively *Focus-on-Form* instruction (FonF) and *Focus-on-FormS* instruction (FonFs). According to Long (1991), who introduced this distinction, FonF instruction "...overtly draws students' attention to linguistic elements as they arise incidentally in lessons whose overriding focus is on meaning or communication" (p. 45–46), whereas FonFS "always entails isolation or extraction of linguistic features from context or from communicative activity" (Doughty & Williams 1998a: 3; see also Norris & Ortega 2000: 437–439; Ellis 2001, and the critical discussion by Sheen, this volume).

There are other dimensions along which FFI can vary which cut across the implicit-explicit distinction, such as whether the instruction proceeds deductively or inductively (cf. De Coo 1996; Hendrix, Housen & Pierrard 2002) or whether it is oriented towards the input or towards the learners' own output (cf. Van Patten 1996).

The last set of moderating factors pertains to the particular language feature targeted for instruction. Some language features may be more amenable to instruction (or certain types of instruction) than others but it is unclear what the relevant constraints are. Some researchers have character-

ized the contrast in terms of the traditional domains of linguistic analysis, i.e. whether the target feature pertains to phonology, lexis, morphology, syntax, discourse-pragmatics or sociolinguistics (e.g. N. Ellis 1993; Ellis 2002; Hendrix, Housen & Pierrard 2002; House 1996; Kasper & Rose 2002; Laufer, this volume). But even within one linguistic domain, some features may be more 'instructionable' than others. Possible mediating factors include the perceptual salience of the target feature in the input, its functional or semantic transparency, its communicative load, its relative (typological, linguistic or cognitive) markedness, all of which can be seen as contributing to the relative *complexity* of the target feature (see Hulstijn & De Graaff 1994 and Doughty & Williams 1998b, for useful starting classifications of these factors).

Closely related to, but ultimately independent from the complexity of the target feature is the complexity of the description or explanation of a particular target feature. Metalinguistic rules and pedagogical descriptions can differ in clarity, intelligibility and processability so that a given target feature can be explained in both simple and elaborated terms (Housen, Van Daele & Pierrard, this volume).

### 3.3. Summary

If we consider the evidence from contributions to this volume and from the wider body of empirical ISLA research, there seem to be two key issues which are commanding the attention of researchers:

- (1) How does instruction affect (a) the basic dimensions of SLA (its route, rate and end-state), (b) the basic components of SLA (exposure, learning propensity, internal processes and mechanisms), (c) the major processes of SLA (the internalisation, modification and consolidation of L2 features and knowledge) and (d) the different types of knowledge which L2 learners develop (implicit vs. explicit; declarative vs. procedural)?
- (2) How are the effects and the success of instruction affected by (a) the type of learner, (b) the type of instruction and (c) the type of feature being instructed?

These are first and foremost empirical questions for SLA research to answer. It is the task of a comprehensive theory of SLA to account for the answers found.

#### 4. Organisation of the volume

The study of ISLA is beset with the difficulties inherent to the complexity of the phenomenon under investigation. It is our conviction that no study of a phenomenon as variegated as ISLA can be conducted satisfactorily within the confines of one theoretical framework or one methodological approach, which accounts for the variety of approaches included in this volume.

The interest in ISLA in this volume is theoretical rather than applied, and acquisitional rather than pedagogical. The studies included in this book exemplify what Ellis (1992) has called ‘the linguistic or psycholinguistic approach to instructed SLA’. They take as their starting point either a theory of language or a theory of SLA and seek to test hypotheses based on the theory using data obtained from instructed L2 learners. For some of the studies, the principal aim is theory construction and the classroom serves primarily as a convenient setting in which to carry out empirical work. Examples of this line of research are the chapters by Gor & Chernigovskaya, Menzel, Ranta and Temple. These authors make use of the high(er) degree of control afforded by the instructional setting to investigate variables put forward by theories in Psycholinguistics, Linguistics or Psychology. In addition to contributing to theory building, other chapters are also explicitly concerned with improving instructional efficiency, believing that educational progress is best ensured if it is research-led and if the research is based on a strong theory. This line is represented by the contributions of, among others, Griggs, Järvinen, and Sheen.

The book is divided into four sections.

Section 1, *Investigating cognitive and processing mechanisms in Instructed SLA*, has five chapters which all draw on theories of speech processing and language competence to investigate cognitive, psycholinguistic and linguistic mechanisms in L2 learning and processing by instructed learners.

In the first chapter of Section 1, Liz Temple investigates the relationship between different types of language knowledge (implicit and explicit knowledge, declarative and procedural knowledge) at work in speech planning and production by analyzing the fluency with which L2 learners produce different linguistic structures (verb phrases vs. noun phrases; dependent vs. independent clauses). Her subjects are eleven English-speaking university learners of French as a foreign language who were recorded at the beginning and end of a semester course in French. Fluency is operationalized as

the degree and distribution of hesitation markers (pauses, incomplete words, repeats, repairs) in the learners' utterances. Temple's analyses indicate that these formally instructed learners significantly gain in fluency over a relatively short period of time. Temple discusses her findings within the framework of Levelt's (1989) speech processing model and Paradis' (1994) distinction between implicit and explicit L2 knowledge. She attributes the observed gain in fluency not to a simple speed-up of speech but rather, to a move to higher-level clausal planning and the development of automatic procedures, resulting in a less hesitant production of grammatical phrases.

In Chapter 3, Barbara Menzel reports on two experimental studies investigating how classroom L2 learners extract grammatical regularities from classroom input. In both studies adult Japanese learners of L2 German were asked to indicate the grammatical gender of German nouns appearing on a computer screen by naming the appropriate nominative definite article as quickly and as accurately as possible. In the second study, the validity of the formal gender indicators on the nouns was systematically varied in order to investigate how these adult learners tackled the problem of conflicting gender cues. In each study the assignment latencies and accuracy rates of the subjects' responses were measured. The results of the two experiments are matched to the frequencies of the different gender indicators in the learners' instructional input. Menzel checks her results against the predictions of the Competition Model (Bates & MacWhinney 1989) and the Network Model (Bybee 1995). She sees her results as confirming Bybee's Network Model according to which the type frequency rather than the token frequency is the decisive factor for the productivity of a morphological pattern.

In Chapter 4 Leila Ranta sets out to test the hypothesis that the psychological trait of analytic ability is associated more with grammatical accuracy than with fluency in speaking. To this end, she analyses the grammatical development and fluency of the English oral production of Francophone children in an intensive ESL program in Quebec which promotes oral interpersonal communication skills in the L2 rather than formal accuracy and academic proficiency. Fluency is assessed using six measures including speech rate, pausing, and self-repair variables. Learners are further classified according to their stage of interlanguage development for question forms and for possessive determiners. Two groups of learners with different levels of analytic ability are selected based on their performance in an L1 and an L2 metalinguistic task. The analytic learners are found to outperform the less analytic learners with respect to the production of possessive



determiners; the difference between the two groups for the question forms is not significant though the trend points in the same direction. In contrast, few differences are noted between the groups with respect to fluency. According to Ranta these results suggest that the communication-oriented nature of the instruction in these intensive ESL programs leads to relatively uniform levels of fluency among learners with different analytic abilities, but that the absence of form-focused instruction may disadvantage less analytic learners in terms of their grammatical development.

Chapter 5 by Kira Gor and Tatiana Chernigovskay investigates whether explicit explanation and controlled practice of grammar rules in a communicative foreign language classroom has an effect not only on learners' performance in the L2 but also on their underlying representation and processing of the target features. They explore this issue in a series of experiments on the acquisition and processing of the complex system of verbal morphology in Russian. The results from fifteen American learners of L2 Russian are compared with baseline data from Russian native speakers and interpreted in the light of the debate between single *versus* dual mechanism models of morphological processing. The results show that not only can the L2 learners successfully apply explicitly taught pedagogical rules in their processing of complex verbal morphology but that structured exposure to the target system also leads them to develop native-like processing strategies. The L2 learners rely on their knowledge of statistical probabilities (type frequencies of the verb classes) to identify the default pattern and generalize this to other verb classes. Gor and Chernigovskay attribute the observed differences between L2 and native processing to differences in type frequency between the L2 learners' input and the input frequencies in native Russian. These findings demonstrate the importance of the statistical characteristics of classroom input for the internalization of the target grammar and for the development of native-like processing strategies in L2 learners.

The eleven chapters in Sections 2 and 3 all deal with the role and effects of instruction on aspects of L2 learning or L2 performance. The five chapters in Section 2 deal with form-focused instruction, while the six chapters in Section 3 are concerned with the role of various types of interaction and communication-focused instruction in promoting L2 development.

Section 2, *Investigating the role and effects of Form-Focused Instruction*, opens with a chapter by Ahlem Ammar and Patsy Lightbown, the first in a series of four which directly analyze the role of *explicit* Form-Focused Instruction and *Focus-on-FormS* instruction on the acquisition and mastery



of specific grammatical target structures. The target structure under investigation in Ammar and Lightbown's study is a classic in L2 research: relative clauses. Building on previous work by Keenan & Comrie (1977), Gass (1982) and Hamilton (1994), the authors set out to test the Implicational Generalization Hypothesis (IGH) according to which instruction of marked items in a hierarchy of implicationally related grammatical structures not only promotes the acquisition of the target relative structures but also the acquisition of less marked relative structures in the hierarchy that were not targeted by the instruction. In addition, Ammar and Lightbown investigate the possibility that the generalization to other relative clause types is not strictly unidirectional so that some subjects may generalize what they were taught not only to less marked relative structures but also to more marked structures that are not implicated by the relative clause type that was taught. In order to test the various hypotheses, three groups of Arabic-speaking learners of L2 English in Tunisia were taught relative clause types exemplifying different levels of markedness. The instructional treatment can be characterized as explicit, Focus-on-FormS as it involved explicit rule presentation and controlled practice of the target structure. The three experimental groups are all found to show better command of relativization on post-test measures than an uninstructed control group. Ammar and Lightbown's results also support the IGH insofar as the subjects generalize their knowledge of relativization to relative clause structures implicated by those that were taught. However, subjects are also found to generalize their newly acquired knowledge to relative clause structures that are more marked and not implicated by those explicitly taught. This surprising finding casts further doubt on the claimed unidirectionality of the Implicational Generalization Hypothesis.

Chapter 7, by Nina Spada, Patsy Lightbown & Joanna White, deals with the importance of form-meaning mappings in explicit form-focused instruction. The impetus of the study is the observation that L2 learners often have great difficulty acquiring L2 features which show a misleading similarity to L1 features. The authors hypothesize that in such cases learners will benefit most from instruction which is not only explicit with regard to the L2, but which also explicitly draws their attention to the contrast between L1 and L2. To investigate this hypothesis, Canadian francophone learners of L2-English in two intact classes received explicit form-focused instruction on the possessive determiners *his* and *her*, and those in two others classes received instruction on question formation. The subjects who were taught possessive determiners outperform control subjects on a variety of post-test measures assessing knowledge of this feature. In contrast, the subjects who

were taught question formation do not outperform their uninstructed control peers. Spada, Lightbown and White attribute these mixed results to differences in the communicative force of the two target structures: failure to use or understand possessive determiners correctly is more likely to cause communication problems than is failure to use or understand questions without the required inversion. More generally, the finding that instruction differentially affects the two target structures in this study points to the influence of form-meaning mappings on the effectiveness of explicit form-focused instruction.

In chapter 8, Alex Housen, Michel Pierrard & Siska Van Daele present a quasi-experimental study which explores the effects of explicit instruction on the learning of grammatical structures in L2 French and the impact of the target structure on the effectiveness of the explicit instruction. The specific attribute of the target structure investigated is its *complexity*, here defined in terms of Givón's (1995) notion of *functional markedness*. In the experiment, two groups of Dutch-speaking learners of French as a foreign language in Belgium were each taught a different grammatical target structure: one group received explicit instruction on a complex structure (the French passive), while the other group was taught a simple(r) structure (French negation). A control group received no instruction on either target structure. Differences between the groups in knowledge and use of the target structures before and after instruction were measured using three tasks varying in the degree of planning and activation of explicit vs. implicit language knowledge: a grammaticality judgement task, a controlled written production task, and an unplanned oral production task. The results indicate that the instructed learners showed significantly better mastery of the target structures on posttest measures than the uninstructed control group, particularly in their unplanned oral production. This result suggests that explicit instruction not only affects the development of explicit language knowledge, as previous studies have shown, but also of implicit knowledge. A second series of analyses shows that the subjects who received instruction on the complex structure made systematically more progress after instruction than the subjects who were taught the simpler structure. This difference, while not statistically significant, is nevertheless suggestive of a relationship between the complexity of the target structure and the effectiveness of explicit instruction.

The studies in chapters 6 to 8 all show clear beneficial effects of explicit, Focus-on-FormS instruction on the learning and use of specific grammatical target structures, thereby challenging current beliefs to the contrary that "formal instruction [...] often fails to teach learners specific linguistic fea-

tures” (Ellis 1994: 107; see also Krashen 1992, 1993, 1994; Long 1991). The effectiveness of explicit, Focus-on-FormS instruction is also demonstrated in Chapter 9, by Ronald Sheen. In the first part of his article, Sheen critically examines basic tenets in the field of applied linguistics and second language acquisition (AL–SLA) in terms of the alleged new teaching approaches that have emerged in recent decades to account for what Sheen calls “the apparent syndrome of failure to bring about identifiable improvement in classroom second-foreign language learning”. In the light of this syndrome, Sheen goes on to examine the current advance of the Focus-on-Form approach to grammar instruction, as advocated by *inter alia* Long (1991), at the expense of a Focus-on-formS. Sheen contends that the current advocacy of Focus-on-Form displays the same shortcomings as previous ineffective advocacies (e.g. Audiolingualism, the Natural Approach and other strong communication-oriented approaches to L2 teaching). Sheen then experimentally compares the effects of a Focus-on-FormS treatment with a Focus-on-Form treatment on the learning and mastery of English interrogatives and adverb placement. Subjects are two groups of Francophone 6<sup>th</sup> grade elementary students of English in Quebec. The Focus-on-FormS group outperforms the Focus-on-Form group on a series of post-test measures involving aural comprehension, grammaticality judgements and controlled oral production of the target features, which points to the greater effectiveness of a Focus-on-FormS in helping students to learn and use grammatical features. On the basis of findings such as these, and “the less-than-compelling findings in favour of a focus on form”, Sheen advocates against L2 instructional reforms which exclude a Focus-on-FormS from the range of teaching strategies available to the L2 instructor.

In the last chapter of Section 2, Batia Laufer takes the claims made by proponents of form-focused instruction for grammar learning and extends them to the domain of vocabulary learning. She starts by considering whether the Default Hypothesis of vocabulary acquisition, originally proposed for first language acquisition, can be extended to instructed SLA. The Default Hypothesis claims that native speakers acquire most of their words through exposure to (esp. written) input, rather than through instruction. Laufer critically reviews five assumptions underlying the Default Hypothesis, concluding that none of them can be taken for granted in an instructional L2 learning context. She also rejects the basic assumption that people possess large vocabularies, too large to be accounted for by form-focused instruction, contending that foreign language learners typically have small vocabularies, sometimes in spite of many hours of instruction. She then proposes an alternative hypothesis, the Planned Lexical Instruction

hypothesis, arguing that, rather than being the result of implicitly ‘picking up’ words from reading, L2 vocabulary knowledge in foreign language contexts mainly results from ‘word focussed classroom instruction’, particularly instruction which “ensures noticing, provides correct lexical information, and creates opportunities for forming and expanding knowledge through a variety of word focused activities”. As such, the Planned Lexical Instruction Hypothesis is in line with both Focus-on-Form and Focus-on-Forms approaches to L2 instruction.

Section 3, *Investigating the role of Interaction and Communication-Focused Instruction*, opens with the chapter by Katja Lochtmann. Lochtmann explores the role of ‘negotiation of form’ by means of negative feedback in foreign language teaching through a comparative discourse-analytic study of oral error correction in an analytic foreign language teaching setting (German as a foreign language in Flanders, Belgium). She investigates the frequency and distribution of six corrective feedback types (explicit corrections, recasts, clarification requests, metalinguistic feedback, elicitations and repetitions) and types of learner uptake following the feedback and compares these to the findings of previous research on corrective feedback in more experiential language teaching settings such as Canadian French immersion. Lochtmann’s findings indicate that teachers in the analytic classrooms are more likely to push learners to correct themselves by means of elicitation and metalinguistic feedback whereas the teachers in the experiential settings more frequently take recourse to recasts. Overall, the learners in the analytic setting receive significantly more negative feedback than those in the experiential setting but in both settings more than half of the feedback fails to achieve its goal, thus casting doubt on the effectiveness of this instructional technique.

In Chapter 12, Kuiken and Vedder report on a study of the role of noticing induced by focus-on-form activities, interaction and metacognition in the learning of English passive structures by Dutch high school students. Two groups of students performed two dictogloss tasks which required them to reconstruct in writing a text previously read to them. The experimental group had to reconstruct the original text in small groups (with the possibility of interaction) while the learners in the control group had to reconstruct the text individually (without the possibility of interaction). Kuiken and Vedder hypothesize that the experimental group will score better on two post-tests measuring recognition of passives and will produce more and more accurate passives in their reconstructed texts than the control group. Neither hypothesis is confirmed by the results of a quantitative

analysis of the data. However, a qualitative analysis of the metatalk that occurred during the interactions between the experimental subjects in the reconstruction phase of the dictogloss tasks, reveals that metacognitive awareness and noticing of the target structures has effectively taken place.

In Chapter 13, Maria del Pilar García Mayo considers whether L2 learners' needs in communication-oriented instructional settings can be effectively addressed through interaction other than negotiation of meaning. To this end she analyses the interactional strategies used by fourteen advanced Spanish university students of English engaged in two dyadic communication tasks with other advanced L2 learners and with native speakers. Special attention is paid to the learners' spontaneous use of repair (self- and other) and collaborative discourse (completion strategies), two types of strategies which provide evidence that learners are focussing on grammatical, phonological and lexical form. García Mayo's analyses suggest that these advanced learners employ both strategies to modify their interactional discourse at the lexical and phrasal level. These two strategies prove insufficient however, both when it comes to meet the learners' need for positive and negative input and to enable them to produce the modified output required for morpho-syntactic growth. García Mayo concludes that communicative language teaching alone is unable to promote advanced levels of grammatical accuracy in a foreign language context. Therefore it needs to be supplemented by focus-on-form activities (e.g. dictogloss tasks, jig-saw strip, story narration), targetting specific problematic form-meaning relationships above the word and phrasal level.

In Chapter 14, Peter Griggs reports on a study of French-speaking learners of English in France to show that metalinguistic activity during communicative L2 production (i.e. negotiation of meaning) in an instructional setting has direct, positive effects on learners' interlanguage development. Learners who show a high rate of metalinguistic activity during communicative task performance are found to make significantly more progress in terms of accuracy, and slightly more progress in terms of fluency, than learners with low rates of metalinguistic activity. This is illustrated by a detailed analysis of the development of past tense and the metalinguistic behaviour (hesitations, repetitions, self-repairs, translations to and from the L1) of one French intermediate learner of English. The results are discussed within the framework of Anderson's (1993) cognitive model and the distinction between declarative and procedural knowledge. Griggs argues that the main contribution of communication tasks to the development of L2 production skills does not so much reside in the communicative processes which they trigger but, rather, in the ways learners use these processes to

reconcile their own internal interlanguage rules and the external target language rules and to adjust these rules to their own communicative needs and target language requirements. This mechanism, Griggs argues, is reminiscent of Anderson's notion of *cognitive tuning*.

In Chapter 15 Heini-Marja Järvinen presents a quasi-experimental study of the syntactic development of young instructed English L2 learners in Finland using L2 English as a medium of instruction according to the principles of *Content and Language Integrated Learning* (CLIL). Järvinen first discusses the theoretical underpinnings of CLIL, arguing that it favours implicit language learning. Drawing on VanPatten (1996), she maintains that a CLIL experience leads learners to attend primarily to the meaning of the linguistic input they receive. As a result of extended exposure to the rich linguistic input offered by a CLIL context, learners are able to implicitly extract grammatical rules and regularities and to integrate them in their subconscious language knowledge store so that they become available for automatic processing. To test this hypothesis, Järvinen examines the development of English relativization from the first grade to the fifth grade in a primary school CLIL program. Elicited imitations and grammaticality judgements are used as measures of interlanguage development of the target feature. Subjects are ninety CLIL pupils and some fifty peer controls from mainstream English foreign language classes in Finland. Quantitative and qualitative analyses of the data show that the development of the syntactic target structures within the experimental CLIL group is faster and more versatile than that of the peer control group. Järvinen's results not only corroborate the findings of a small yet growing body of research that CLIL is a useful L2 learning-teaching method but they also underline the importance of attending to meaning and the role of implicit learning processes in the L2 acquisition of complex grammatical knowledge.

In the last chapter of Section 3, Tsuyoshi Kida discusses the socio-psychological dimension of the acquisition of communicative competence in an instructed L2 learning context. Kida first expounds how this socio-psychological dimension is closely tied to the workings of declarative and procedural memory. He then shows how the experience of interaction is a significant didactic problem for both the implicit and explicit teaching of communicative competence. The challenge for classroom participants is to engage in meaningful exchanges while at the same time take into account the linguistic, personal and cultural backgrounds of the other participants. These issues are further elucidated in a series of qualitative and quantitative analyses of the discourse patterns that occurred in one French L2 class attended by seven adult immigrants to France from various L1



backgrounds. Kida's analyses suggest that, firstly, the language teacher contributes to the development of learners' discourse competence through the pragmatic and interactional context she creates through her discourse and, secondly, that this context plays a major role in shaping the learners' discourse, both qualitatively in terms of the discourse patterns adopted and quantitatively, in terms of the frequency with which these discourse patterns occur. The chapter concludes with a consideration of the conditions under which L2 learners could benefit from classroom interactions with a view to improving their discourse competence.

Section 4, *Comparing Instructed and Naturalistic L2 Learning Contexts*, has two chapters which both examine the impact of the global acquisition context (instructed, uninstructed or mixed) on selected aspects of learners' L2 competence.

Chapter 17, by Martin Howard, represents the growing body of research on the *Study Abroad* experience (cf. Freed 1995, 1998). Howard adopts the variationist approach and the analytic framework of lexical semantics in a detailed comparative analysis of the grammatical expression of past time in the French oral interlanguage of three groups of advanced Irish university learners of French. The aim of the study is to compare the effect of naturalistic contact with the target language (as experienced during a year-long stay in France) relative to the effect of prolonged formal classroom instruction in the home country. Detailed quantitative and qualitative analyses reveal differences and similarities between the various groups, which demonstrate the complexity of identifying, in absolute terms, whether study abroad benefits grammatical development in an L2 more than classroom instruction. Howard's results suggest that study abroad and formal classroom instruction do not have radically different effects on the learners' underlying grammatical patterns and representations, i.e. their grammatical competence. However, when it comes to the learners' ability to actually use the target language grammatical morphology in real time contexts, i.e. their grammatical skills, the study abroad experience is shown to be more beneficial than the prolonged formal instruction. This is reflected in the increased use and higher level of accuracy of use of the various French past time markers by the study abroad learners. The finding that the differential effect of study abroad and formal classroom instruction merely involves a change in the learners' level of use of temporal-aspectual markers but not in how they use these markers points to the effect of semantic and pragmatic universals in the acquisition and use of the tense and aspect system, irrespective of learning context.

In the final chapter, Jean-Marc Dewaele surveys the impact of the language learning context (instructed, naturalistic and mixed) on the development of sociolinguistic, sociopragmatic and sociocultural aspects of communicative competence in a second language. Previous research suggests that mixed contact (instructioned + naturalistic) leads to higher levels of these competencies compared to purely classroom-based instruction. However, it is not clear if this is also true when intensity and amount of exposure are held constant. Dewaele further explores this issue by examining the impact of the learning context on two variables reflecting sociopragmatic competence, namely perception of locutionary force of swearwords and self-reported language choice for swearing. Dewaele collected data for second, third, fourth and fifth languages (defined in terms of timing and sequence of acquisition) from more than 1,000 multilinguals through a web questionnaire containing both closed Likert-type questions and open questions. Statistical analyses of the answers indicate that, as expected, learning context/type of contact indeed has a significant effect on the investigated aspects of communicative competence, with the instructed learners scoring almost consistently lower than both the mixed learners and the naturalistic learners on self-reported use and perceived locutionary force of swearwords. However, further analysis shows that frequency of use of the target language and age of onset of acquisition have an even stronger effect and make better predictors of learners self-reported sociopragmatic competence in consecutive non-native languages. These findings lead Dewaele to conclude that “whatever type of language instruction/contact one opts for, the results suggest that the younger one starts it, the better the end result will be”.

## **5. Conclusion**

As one early reviewer pointed out (Harley 1988), investigating instructed second language acquisition is a fascinating, but daunting, task: fascinating, because it addresses fundamental issues in language acquisition and language knowledge and how they are affected by varying contextual conditions; daunting, because of the complex, multidimensional nature of both SLA and instruction, and the interacting effects of social, linguistic, cognitive and personality factors. The chapters in this volume illustrate not only the productivity but also the diversity of current research on instructed second language acquisition, diversity in terms of issues investigated, theoretical orientation and methodology. The heterogeneity of topics and methodology does not contradict the homogeneity of purpose though, which is to contribute to theory construction and to inform L2 education.



## Notes

1. Except where stated otherwise, the terms 'acquisition' and 'learning' are used interchangeably in this chapter.
2. Consciousness can also be equated with intentionality, although this is not considered necessary for language learning as non-intentional or incidental learning is also possible (Schmidt 1995; Ellis 1997).

## References

- Anderson, J.  
1993 *Rules of the Mind*, Hillsdale, NJ: Lawrence Erlbaum.
- Bates, E. and B. MacWhinney (eds.)  
1989 *The Crosslinguistic Study of Sentence Processing*. New York: Cambridge University Press.
- Bardovi-Harlig, K.  
2000 *Tense and Aspect in Second Language Acquisition: Form, Meaning, and Use*. Oxford: Blackwell.
- Bialystok, E.  
1994 Analysis and control in the development of second language proficiency. *Studies in Second Language Acquisition* 16: 157–168.
- Bybee, J.  
1995 Regular morphology and the lexicon. *Language and Cognitive Processes* 10: 425–455.
- Chaudron, C.  
1988 *Second Language Classrooms: Research on Teaching and Learning*. Cambridge: Cambridge University Press.
- Decoo, W.  
1996 The induction-deduction opposition: ambiguities and complexities of the didactic reality. *International Review of Applied Linguistics*, 34: 95–118.
- De Graaff, R.  
1997 *Differential Effects of Explicit Instruction on Second Language Acquisition*. Den Haag: Holland International Graphics.
- DeKeyser, R.  
1995 Learning second language grammar rules: An experiment with a miniature linguistic system. *Studies in Second Language Acquisition*, 17 (3): 379–410.
- Doughty, C. and J. Williams  
1998a Issues and Terminology. In *Focus on Form in Classroom Second Language Acquisition*, C. Doughty and J. Williams (eds.), 1–11. Cambridge: Cambridge University Press.

- Doughty, C. and J. Williams  
 1998b Pedagogical choices in focus on form. In *Focus on Form in Classroom Second Language Acquisition*, C. Doughty and J. Williams (eds.), 197–261. Cambridge: Cambridge University Press.
- Eckman, F., L. Bell and D. Nelson  
 1988 On the generalisation of relative clause instruction in the acquisition of English as a Second Language. *Applied Linguistics* 9 (1): 1–20.
- Ellis, R.  
 1984 *Classroom Second Language Development*. Oxford: Pergamon.  
 1985 *Understanding Second Language Acquisition*. Oxford: Oxford University Press.  
 1990 *Instructed Second Language Acquisition*, Oxford: Basil Blackwell.  
 1992 Learning to communicate in the classroom. *Studies in Second Language Acquisition* 14: 1–23.  
 1994 *The Study of Second Language Acquisition*. Oxford: Oxford University Press.  
 1997 Explicit knowledge and second language pedagogy. In *Encyclopedia of Language and Education, Volume 6: Knowledge about Language*, L. van Lier and D. Corson (eds), 109–118. Amsterdam: Kluwer Academic Publishers.  
 1999 *Learning a Second Language Through Interaction*. Amsterdam: John Benjamins.  
 2001 Investigating form-focussed instruction. *Language Learning* 51: supplement 1. *Form-focussed instruction and Second Language Learning*, 1–46.  
 2002 Does form-focused instruction affect the acquisition of implicit knowledge?: A review of the research. *Studies in Second Language Acquisition*, 24 (2): 223–236.
- Felix, S.  
 1981 The effect of formal instruction on language acquisition. *Language Learning*, 31: 87–112
- Freed, B.  
 1995 Language learning and study abroad. In *Second Language Acquisition in a Study Abroad Context*, B. Freed (ed.), 3–33. Amsterdam/Philadelphia: Benjamins.  
 1998 An overview of issues and research in language learning in a study abroad setting. In B. Freed (ed.), *Frontiers: The Interdisciplinary Journal of Study Abroad. Special Issue: Language Learning in a Study Abroad Context IV*: 31–60.
- Gass, S.  
 1982 From theory to practice. In *On TESOL 81*, M. Hines and W. Rutherford (eds.), 129–139. Washington DC: TESOL.  
 1989 Language universals and second language acquisition. *Language Learning* 39: 497–534.

- Givón, T.  
 1995 *Functionalism and Grammar*. Amsterdam/Philadelphia: Benjamins.
- Hamilton, R.  
 1994 Is implicational generalisation unidirectional and maximal? Evidence from relativisation instruction in a second language. *Language Learning* 44: 123–157.
- Harley, B.  
 1988 Effects of instruction on SLA: Issues and evidence, *Annual Review of Applied Linguistics* 9: 165–178.
- Hendrix, L., Housen, A. and M. Pierrard  
 2002 Impact et efficacité de différents types d’enseignement axés sur la forme pour l’acquisition d’une langue étrangère. *Marges Linguistiques* 4: 23–49.
- House, J.  
 1996 Developing pragmatic fluency in English as a Foreign Language: Routines and metapragmatic awareness. *Studies in Second Language Acquisition* 18 (2): 225–252.
- Hulstijn, J. and R. de Graaff  
 1994 Under what conditions does explicit knowledge of a second language facilitate the acquisition of implicit knowledge? A research proposal. In *Consciousness in Second Language Learning*, J. Hulstijn and R. Schmidt (eds.), *AILA Review* 11: 97–109.
- Johnson, K.  
 1996 *Language Teaching and Skill Learning*. Oxford: Blackwell.
- Kasper, G. and K. Rose  
 2002 *Pragmatic Development in a Second Language*. Oxford: Blackwell.
- Keenan, E. and B. Comrie  
 1977 Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8: 63–99.
- Klein, W.  
 1986 *Second Language Acquisition*. Cambridge: Cambridge University Press.
- Krashen, S.  
 1981 *Second Language Acquisition and Second Language Learning*. Pergamon: Oxford.  
 1985 *The Input Hypothesis: Issues and Implications*. London: Longman.  
 1992 Formal grammar instruction: Another educator comments... *TESOL Quarterly*, 26 (2): 409–411.  
 1993 The effect of formal grammar instruction: Still peripheral. *TESOL Quarterly*, 27 (4): 722–725.  
 1994 The input hypothesis and its rivals. In *Implicit and Explicit Learning of Languages*, N. Ellis (ed), 45–77. London: Academic Press.
- Larsen-Freeman, D. and M. Long  
 1991 *An introduction to second language acquisition research*. New York: Longman.

- Levelt, W.  
 1989 *Speaking: From Intention to Articulation*. Cambridge, Mass.: MIT Press.
- Long, M.  
 1983 Does second language instruction make a difference? A review of research. *TESOL Quarterly* 17 (3): 359–82.  
 1991 Focus on Form: A design feature in language teaching methodology. In *Foreign language research in cross-cultural perspective*, K. de Bot, R. Ginsberg, and C. Kramsch (eds.), 39–52. Amsterdam: Benjamins.
- Norris, J. M. and L. Ortega  
 2000 Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning* 50 (3): 417–528.
- McLaughlin, B.  
 1987 *Theories of Second-Language Learning*. London: Edward Arnold.
- Paradis, M.  
 1994 Neurolinguistic aspects of implicit and explicit memory: Implications for bilingualism and SLA. In *Implicit and Explicit Learning of Languages*, N. Ellis (ed.), 393–419. San Diego: Academic Press.
- Pienemann, M.  
 1984 Psychological constraints on the teachability of languages. *Studies in Second Language Acquisition* 6 (2): 18–214.  
 1987a Determining the influence of instruction on L2 speech processing. *Australian Review of Applied Linguistics*, 10: 83–113.  
 1987b Psychological Constraints on the Teachability of Languages. In *First and Second Language Acquisition Processes*, Pfaff, C. (ed.), Cambridge, Mass.: Newbury House.
- Ranta, L.  
 2002 The role of learners language analytic ability in the communicative classroom. In *Individual differences and instructed language learning*, P. Robinson (ed.), 159–80. Amsterdam/Philadelphia: Benjamins.
- Reber, A.  
 1993 *Implicit Learning and Tacit Knowledge*. Oxford: Clarendon Press.
- Robinson, P.  
 1996 *Consciousness, Rules, and Instructed Second Language Acquisition*. New York: Peter Lang.
- Sawyer, M. and L. Ranta  
 2001 Aptitude, individual differences and L2 instruction. In *Cognition and Second Language Instruction*, P. Robinson (ed.), 319–353. Cambridge: Cambridge University Press.

- Schmidt, R.  
1990 The role of consciousness in second language learning. *Applied Linguistics* 11: 129–158.  
1995 Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In R. Schmidt (ed.), *Attention and awareness in foreign language learning*, 1–64. Honolulu: National Foreign Language Resource Center.  
2001 Attention. In *Cognition and second language instruction*, P. Robinson (ed.), 3–32. Cambridge: Cambridge University Press.
- Sharwood Smith, M.  
1993 Input enhancement in instructed SLA: Theoretical bases. *Studies in Second Language Acquisition* 15 (2): 165–179.
- Skehan, P.  
1989 *Individual Differences in Second Language Learning*. London: Edward Arnold.  
2002 Theorising and updating aptitude. In *Individual differences and instructed language learning*, In P. Robinson (ed.), 69–93. Amsterdam/Philadelphia: Benjamins.
- Spada, N.  
1997 Form-focused instruction and second language acquisition: A review of classroom and laboratory research. *Language Teaching* 30: 73–87.
- Swain, M.  
1985 Communicative Competence: some roles of comprehensible input and comprehensible output in its development. In *Input in Second Language Acquisition*, S. Gass and C. Madden (eds.), 235–253. Rowley, Mass.: Newbury House.
- VanPatten, W.  
1990 Theory and research in Second Language Acquisition and Foreign Language Learning: On producers and consumers. In *Second Language Acquisition – Foreign Language Learning*, W. VanPatten & J. Lee (eds.), 17–27. Clevedon: Multilingual Matters.  
1996 *Input Processing and Grammar Instruction in Second Language Acquisition*. Norwood, NJ: Ablex.
- Wong-Fillmore, L.  
1989 Teachability and second language acquisition. In *The Teachability of Language*, M. Rice & R. Schiefelbusch (eds), 311–332. London: Paul H. Brookes.



# **Investigating cognitive and processing mechanisms in instructed SLA**





# **Instructed learners' fluency and implicit/explicit language processes**

*Liz Temple*

This study investigated instructed L2 learners' fluency and its relation to the planning and encoding of spontaneous speech. Eleven adult learners of French were recorded twice, at the beginning and end of a university semester course. Fluency was quantified by measuring temporal variables and improvement was evaluated. Surface structure constituents were examined in order to discover whether certain types of grammatical encoding involved significantly more hesitation. While hesitation could not be linked to any particular structure, it was found that fluent and non-fluent learners performed differently relative to how hesitation was distributed throughout the utterance, fluent learners producing most hesitation at the beginning of clauses. It is suggested that this is the result of fluent learners having developed clausal planning. It is proposed that the type of knowledge accessed in speech production, implicit or explicit, associated with automatic or controlled processes respectively, is one factor accounting for learner fluency. Pedagogical implications are discussed.

## **1. Introduction**

This study reports on an investigation into instructed language learners' fluency in spontaneous speech. In the term fluency I am considering the temporal organisation of speaking: "fluent" describes speech produced at a normal rate and which is not unduly disrupted by pauses and hesitations (even though these are a normal part of speech). This definition of "fluent" is more restricted than the broader notion of fluency in a language which often refers to a general speaking proficiency (see Koponen and Riggensbach 2000 for a discussion of the term "fluency"). Firstly, the study of the temporal organisation of speech, given that the act of speaking occurs "on-line", can lead to an understanding of the nature of language processing. Speech involves the integrated functioning of multiple cognitive and linguistic systems: the demands of thought, attention, memory, along with lexical, semantic, syntactic, morphological, and phonological processing and pragmatic

situation, are integral to the time course of speech production. Secondly, this kind of fluency can be measured: temporal variables such as speech rate, pause rate and articulation time ratio are indicative of the time needed for processing activity. Where processing is constrained in some way, it will require more time, and production will be less fluent. For example, if the task is more demanding, the speaker will hesitate more: native speakers have been shown to pause twice as much in an interpretation task as in a simpler description task (Goldman Eisler 1968). Planning, selection, encoding and monitoring may require time in spontaneous speech; when repeating this speech, both natives and nonnatives improve their fluency as pause time gradually decreases, since many of the time-demanding procedures are eliminated or reduced in the repetition task (Goldman Eisler 1968, Nation 1989). If task difficulty and cognitive activity are fundamental to the temporal organisation of native speech, it is apparent that learner speech will be even more time-constrained.

Studies of language learners' (L2) fluency have investigated differences between natives and learners (Wiese 1984, Temple 2000), or differences over time or period of instruction, in particular, differences before and after a study abroad period (Raupach 1984, 1987, Lennon 1990, Towell et al. 1996, Freed 1998). In this study, learner fluency was assessed over a fairly short period of instruction and the placement of pauses within the utterance was examined. The aim was to discover, firstly, if a gain in fluency was possible within this period, and secondly, whether hesitation is distributed equally throughout the utterance or whether it occurs to a greater extent in specific segments, thus reflecting difficulty in encoding in a certain type of grammatical structure or functional unit. In particular, the verb group was compared to the noun group, since the hypothesis that the verb group was especially a source of uncertainty and hesitation had been suggested by an earlier finding that the learners repaired errors in verb groups more frequently than those in noun groups (Temple 1992). If verbs involve a greater amount of monitoring, this could be due to grammar instruction emphasising verb forms in particular, or an inherent complexity of processing French verb morphology for speakers of English. In either case, we need to verify whether the verb group requires extra time to process.

Thirdly, the degree of fluency of the learner was also taken into account, given that development of fluency might interact with type of syntactic unit, such that a subject's level of fluency might have a bearing on which elements are more or less hesitation prone. It could be, for example, that a fluent learner produces particular parts of the utterance in a more hesitant manner.

In brief, the questions in this study are: (1) Can instructed learners, without study abroad opportunities and over a short period of instruction, improve in fluency? (2) Do structural units of the utterance show differing degrees of hesitation? (3) If so, does this depend on how fluent the learner is? The results of my analysis will be followed by a discussion of the implicit/explicit learning distinction and its relevance to the question of learner fluency in instructional settings.

## **2. Language learner hesitation**

Given the language learner's lower level of competence, one would expect considerably more frequent use of hesitation phenomena than for native speakers, and in particular more hesitation associated with syntactic and lexical encoding difficulties. The following extract of one learner's text illustrates this point.

### 2.1. Learner speech sample

Transcription conventions:

// silent pause between .2 – .5 sec.  
/// silent pause between .5 – 1 sec.  
//// silent pause greater than 1 sec.  
: lengthened syllable  
italics – repetition or incomplete word

(Each line represents a fluent run.)

Je crois que: beaucoup de gen:s avont  
*//euh* ont  
*///euh* les aspiration:s  
*euh///* d'être un acteu:r  
*//euh//* mai:s  
*////* quan:d  
*//mm* a- actuellement c'est  
*c'est c'est* ce n'est pas une  
*//* bonne idée parce que quand  
*////* on a besoin

// d'argent qu'on a besoin de: manger  
 /// alors  
 // ce ce profession ne:  
 n'off- // n'offre pas  
 /// beaucoup

In this extract, silent and filled pauses, along with hesitations such as incomplete words (*a-*, *n'off-*) and repeats (*c'est c'est*, *ce ce*), disrupt the sentence and reveal the speaker's difficulties in formulating his ideas in French. The student appears to be closely monitoring his speech, for example, the verb form is successfully revised in the repair *avont – ont*, and the hesitation in *n'off- n'offre* may reveal his struggle to find the verb ending. Long pauses preceding the conjunctions and connectors *mais*, *quand* and *alors*, provide time for planning, though it is clearly insufficient. Fragmented syntactic units such as *une // bonne idée* are clues to problems of lexical search or morpho-syntactic encoding while the units are in mid-production. The extract contains 17 hesitations in a sentence of 45 words.

## 2.2. Studies in learner fluency

The above example of foreign speech is typical of the lack of fluency in instructed learners. Analyses of temporal variables in learner speech have led to findings of a significantly lower speech rate, a higher pause frequency (or longer mean length of run) and longer duration of silent pauses. Researchers have concluded that this is principally due to a lack of automatization or proceduralisation (e.g. Wiese 1984, Raupach 1987, Temple 1992, Towell et al. 1996). Most of the research has linked an improvement in fluency to a study abroad period: learners who are in foreign language instructional settings are less likely to substantially proceduralise their knowledge until they experience study abroad. "A core feature of development in advanced L2 learners therefore seems to be the conversion of linguistic knowledge already acquired into rapidly-usable on-line 'productions'", this occurring typically after 6 months abroad in a four-year longitudinal study by Towell et al. (1996: 113). Freed (1998) found a significant improvement in perceived fluency for students who had a study-abroad semester compared with those who did not, but only for the less advanced learners of French in her study, i.e. where there was greater capacity for improvement (or for improvement to be perceived). It is not surprising that our belief in the advantages of study abroad is confirmed by

such studies. Indeed, it could be more illuminating to learn if instructed learners can gain in fluency during their course at home. And if fluency is indicative of proceduralisation, can we determine which kind of linguistic computations become proceduralised?

### **3. Method**

The subjects were eleven English speaking adult learners of French, university students at intermediate and advanced level, who had learned French in a formal instruction setting for between one and six years. Their current studies included a 2-hour unit devoted to listening and speaking skills. The students were also following a concurrent core language unit (2 hours for the advanced group, 4 hours for the intermediate group), which included some metalinguistic instruction, as well as study and practice of the four skills. They were recorded in pairs in an interview task towards the beginning and end of a session-length course, about a three month interval; the two tasks are called "pre-test" and "post-test". The interview topics were chosen by the students in areas of personal or general interest (studies, music etc.) and the teacher-researcher was not present, so that the situation would be as natural as possible. The interview task was chosen to elicit texts from the interviewees with minimal interaction, since pauses associated with turn-taking were not pertinent to the study. Twenty-two samples of speech (11 pre- and 11 post-test) of approximately two minutes length were selected and transcribed.

Silent pauses were measured from mingograph traces of the speech and hesitation phenomena were added to the data. The minimum pause length chosen for calculations could be varied by computer and was initially set at 200 msec. in order to capture all but very short pauses. Two temporal variables were calculated by computer. These were: speech rate, the number of syllables per total speaking time in seconds, which is a function of the amount of pausing involved plus the rate of actual articulation; and articulation time ratio, a measure which is independent of articulation rate but which accounts for the amount of silent pause time, and is the percentage of total time spent in actual verbalising. The following phenomena were counted as hesitations: silent and filled pauses, combined pauses (containing both silence and filled pause), incomplete words, repeats, false starts and repairs. Lengthened syllables were not included for purely technical reasons, as they did not occur as separate elements.

Examples of hesitation phenomena:

filled pause – *euh* or *um*  
 incomplete word – *n'off-*  
 repeat – *c'est c'est*  
 false start – *comment avant la-*  
 repair – *ils avont ont*

Taking the clause as a basic unit of planning, the texts were segmented at the level of clause surface structure: constituents were classified in the five major grammatical functions, subject, verb, object, adverbial adjunct and connector (Garman 1990: 136). The following learner sentence is given as an example of this segmentation, where + indicates a constituent boundary:

/// et alors + il + *euh* // a + une // place + pour // son /// voiture  
 CONNECTOR + SUBJ. + VERB + OBJECT + ADJUNCT

In addition, sub-categories of the constituent types were distinguished, based on grammatical (lexical) categories, so that, in particular, noun phrases could be extracted from the constituents in which they occurred and compared with verb phrases.

The degree of fluency of the learner speech samples was quantified on the basis of the articulation-time-ratio measure (the proportion of total time spent in articulating). Three fluency levels were distinguished by dividing the range of scores into three intervals, the highest scores constituting a high-fluency category (6 speech samples), the mid-range scores a mid-fluency category (8) and the lowest scores a low-fluency category (8). The relationship between the segment type and the occurrence of pauses and hesitations in segments was examined statistically, using SPSS and Genlog software packages. A hesitation ratio was calculated for each segment type by dividing the number of hesitations by the total number of elements i.e. words and hesitations, stated as a percentage. It was found that a pause cut-off of 500 milliseconds discriminated better between the samples, so short pauses of less than 500 msec. were excluded. This cut-off point is consistent with some pause studies (e.g. 500 msec. in Goldman-Eisler 1972, 400 msec. in Freed 1998). Results for high-fluency and low-fluency speech samples were compared (now termed fluent and nonfluent).

#### 4. Results

Improvement in speech rate occurred in 8 of the 11 subjects. However, the average gain over the semester was very small, from 2.27 to 2.41 syllables per second. For the subjects whose speech rate improved, the mean rate rose from 1.97 to 2.23 syll./sec., which is an increase of 13%. It was found that the three learners whose speech rate fell over the three month period of instruction had obtained the highest score in the pre-test, while those with the lowest scores in the pre-test made the most gain (increases of 22%, 18% and 33%).

When the hesitation ratios of major constituents were analysed, the overall test for the constituents was significant (Wald chi-square = 9.44). This means that there was a difference in hesitancy across the five major constituents. To discover where the difference lay, comparisons across levels (between constituent types) were made using loglinear analysis. The results were as follows. No individual constituent was significantly more hesitation prone than the others, however, when combinations of constituents were selected, a difference was found between connector plus subject plus verb compared with object plus adjunct (chi-square = 8.63), hesitation being more frequent in the former combination (see Table 1). The combination subject plus verb group and object plus adjunct (chi-square = 5.87) approached significance, the former pair being more hesitation prone than the latter.

Table 1. Hesitation ratios (as %) for major constituents

	connector	subject	verb	object	adjunct
hesit. ratio	22.0	21.6	20.9	17.7	18.4

It is difficult to see any coherence in terms of grammatical structure in these combinations of segments, and we cannot deduce a relationship between grammatical structure and hesitancy from this particular analysis. However there appears to be a relation with position in the clause: the units which generally occur early in the clause contain significantly more hesitation than those which are normally at the end of the clause.

Now, when the fluency rating of the speech samples was also considered, an interaction was found between fluent/nonfluent learner speech and constituent type (see Table 2). That is, the frequency of hesitations depended

not just on the type of constituents but also on whether the learner was fluent or nonfluent. The most significant interaction was found in the case of connectors. Fluent learners were more hesitant in connectors compared to the other constituents combined, whereas for nonfluent learners the pattern was reversed, that is, nonfluent learners hesitated less in connectors compared to the other constituents.

*Table 2.* Comparison of hesitation ratios for segment types in fluent and nonfluent learner speech

	connector	subject	verb	object	adjunct
fluent	23.0	18.4	19.6	12.5	15.2
nonfluent	18.3	24.4	23.2	19.6	23.8

At the phrase level, where comparisons were made between grammatical categories occurring within different constituent types, in particular between noun groups and verb groups, no differences of interest were found for grammatical categories relative to frequency of hesitation. Our hypothesis that the verb group was more hesitation prone than the noun group was not confirmed. This suggests that encoding difficulty is not specific to any one structure across the group of learners.

Thus we have the following findings regarding our three research questions. (1) Most of the learners made some improvement in speech rate over the semester, but this was not due to a simple speeding up of performance. (2) No particular type of structural unit required more hesitation. (3) While it appeared that, overall, learners showed greater hesitancy early in the clause or sentence, this conclusion must be revised when fluent and nonfluent speech samples are compared. Given that fluent speech involved most hesitation in the connector constituent, whereas in nonfluent speech the opposite occurred, we can see that initial hesitation is not a generalised phenomenon, but is specific to fluent learners only. Their speech is characterised by a shift from clause-internal to clause-initial hesitation.



## **5. Discussion**

### **5.1. Placement of pauses**

If we refer to the research on pause placement in native speech, we find that the shift to clause-initial hesitation is a more nativelike pause behaviour. This research has been summarised by Garrett (1982), who described the significance of pauses for planning at different levels in speech production. Three circumstances which lead to the production of hesitation pauses in the utterance are postulated. Firstly, there is long range planning: pauses occur where processing load is temporarily exceeded by the requirements of macro-planning. Secondly, there is the role played by the structural unit: some surface unit, typically the clause, but also constituents within the clause, requires initial planning time. Thirdly, there is what Garrett terms “the coordination of components”. This means that not all elements of a planning unit like the clause are in fact supplied during planning, and further pauses may be determined by a “filling in the slot” procedure (Garrett 1982: 47). It is held that three types of processing decision – conceptual, syntactic and lexical – are made at certain points in the utterance, with the result that pauses and hesitations may occur more frequently in the initial phase of longer multi-clause units, around clause boundaries, and preceding less predictable words. Goldman-Eisler (1972) calculated that 78% of sentence boundaries in a native speaker English corpus were marked by a silent pause greater than 500 msec., compared with 34% of clause boundaries, and 8.5% of between-word positions.

Thus the shift towards clause-initial pauses in fluent L2 speech can be related to planning units, although long-range planning units cannot be separated from clause-length units in my analysis. At low-level fluency there is no evidence of a clausal unit of planning, and the time spent in planning and encoding speech is distributed across the utterance. The connector segment, at the clause boundary, involves less hesitation than the rest of the clause, which is the opposite of what happens in native speech. Then as fluency develops, clausal planning is preferred to phrase planning. Hesitations still occurred in the remainder of the clause, but significantly less than for nonfluent learners. The syntactic units within the clause can be produced more readily when planning time has been allotted at the beginning. The third type of pause in Garrett’s analysis, for slot-filling after a planned unit has been commenced, may also become less frequent when clausal planning is more efficient in fluent learners. Furthermore, it was found that the particular type of grammatical category or functional unit

being encoded was not a factor in the placement of pauses and hesitations. It was not the case that verb groups gave rise to more hesitation than noun groups, as had been hypothesised.

## 5.2. Implicit and explicit knowledge

It can be seen that the development of fluency is not a simple speeding up of speech production, but a qualitative change in the way language is processed. In instructed learners, language processing may be of different kinds according to the type of linguistic knowledge held and the way it is being used by the learners. The distinction made by Paradis (1998) between implicit and metalinguistic knowledge is particularly relevant to the context of formal instruction.

It is [...] important to distinguish between implicit linguistic knowledge and metalinguistic knowledge. The former is acquired incidentally, is stored in the form of procedural know-how without conscious knowledge of its contents, and is used automatically. The latter is learned consciously, is available for conscious recall, and is applied to the production (and comprehension) of language in a controlled manner. Implicit linguistic competence is acquired through interaction with speakers of the language in situational contexts. Metalinguistic knowledge is usually learned through formal instruction.

(Paradis 1998: 428)

Metalinguistic knowledge is explicit knowledge, the kind of knowledge that is acquired intentionally and that we are consciously aware of and can verbalise once it is acquired. Implicit linguistic knowledge is our internalised knowledge of a language. We acquire implicit knowledge incidentally and apply it without consciously being aware of it. Distinct memory systems support the two modes of learning: declarative (or semantic) memory is associated with explicit knowledge, procedural memory with implicit knowledge. The implicit-explicit distinction raises key issues in the area of second language acquisition and has been much debated in the literature (see Ellis 1994). Learners receiving formal grammar instruction, as was the case in this investigation, have clearly acquired a considerable amount of explicit knowledge. What are the implications for their fluency? Central to the question is the kind of processes involved in using implicit or explicit knowledge, referred to by Paradis above as “automatic” or “controlled”.

### 5.3. Automatic and controlled processes

This distinction was first made in relation to skill learning. A controlled process is described as the temporary activation of a sequence of nodes, utilising short term memory store: at first effort or attention is required for the processing of novel or inconsistent information. After practice or training, a relatively permanent set of associations is established, and stored in long term memory. When the particular input (set of conditions) of a learned sequence is present, an automatic process is activated without the necessity of our attentional control (Shiffrin and Schneider 1977). A complex cognitive activity such as speaking will require a mix of controlled and automatic processes. Segalowitz (2000) highlights the importance of the nature of the balance between the processes and distinguishes cognitive fluency from performance fluency. Cognitive fluency refers to the efficiency of processing achieved by a particular balance, with an improvement in fluency involving a change in the balance, a shift, for example, to greater use of automatic processes. Performance fluency, on the other hand, refers to speed, fluidity and accuracy in performance and does not always reflect cognitive fluency. Increasing performance fluency may be due to a simple speed-up of processes, without any accompanying change in the distribution of automatic and controlled processes.

Now, the more L2 learners apply explicit knowledge, the more they make use of attentional resources, and require time to use the knowledge in a conscious non-automatic manner, resulting in nonfluent speech. Following Segalowitz, fluent L2 learners could become more practised at using controlled processes for applying explicit knowledge, and could be simply revealing performance fluency. However, when greater fluency is not the result of an overall speed-up of speech but is characterised by a shift to clausal planning, as we have found, it is due to a change in cognitive fluency, in the balance between automatic and controlled processes. In other words, there is a restructuring of the underlying mechanisms of speaking the foreign language. An account of the processing components in speech production will be given to examine the likely focus of restructuring.

### 5.4. A speech production model

The implicit and explicit knowledge activated and the balance of controlled and automatic processes used in speaking are coordinated in the stages of processing which underlie speech production. Levelt's (1989) first-language

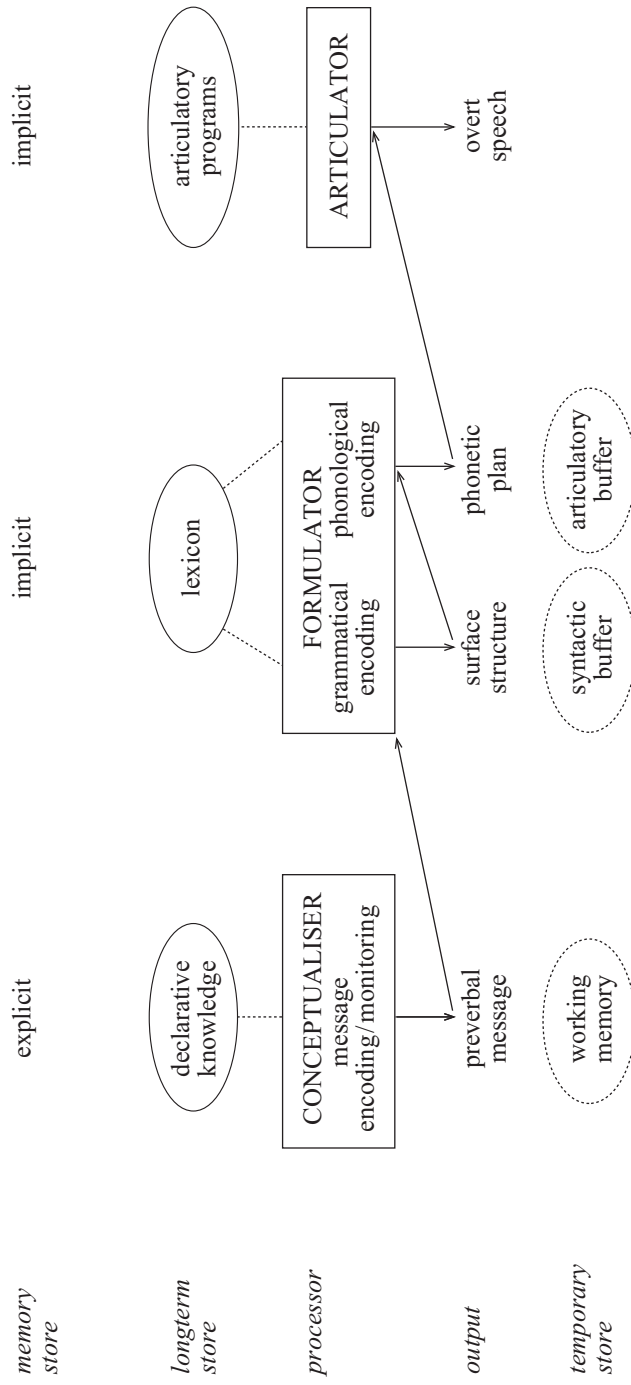


Figure 1. L1 Speech Production Model (after Levelt 1989)

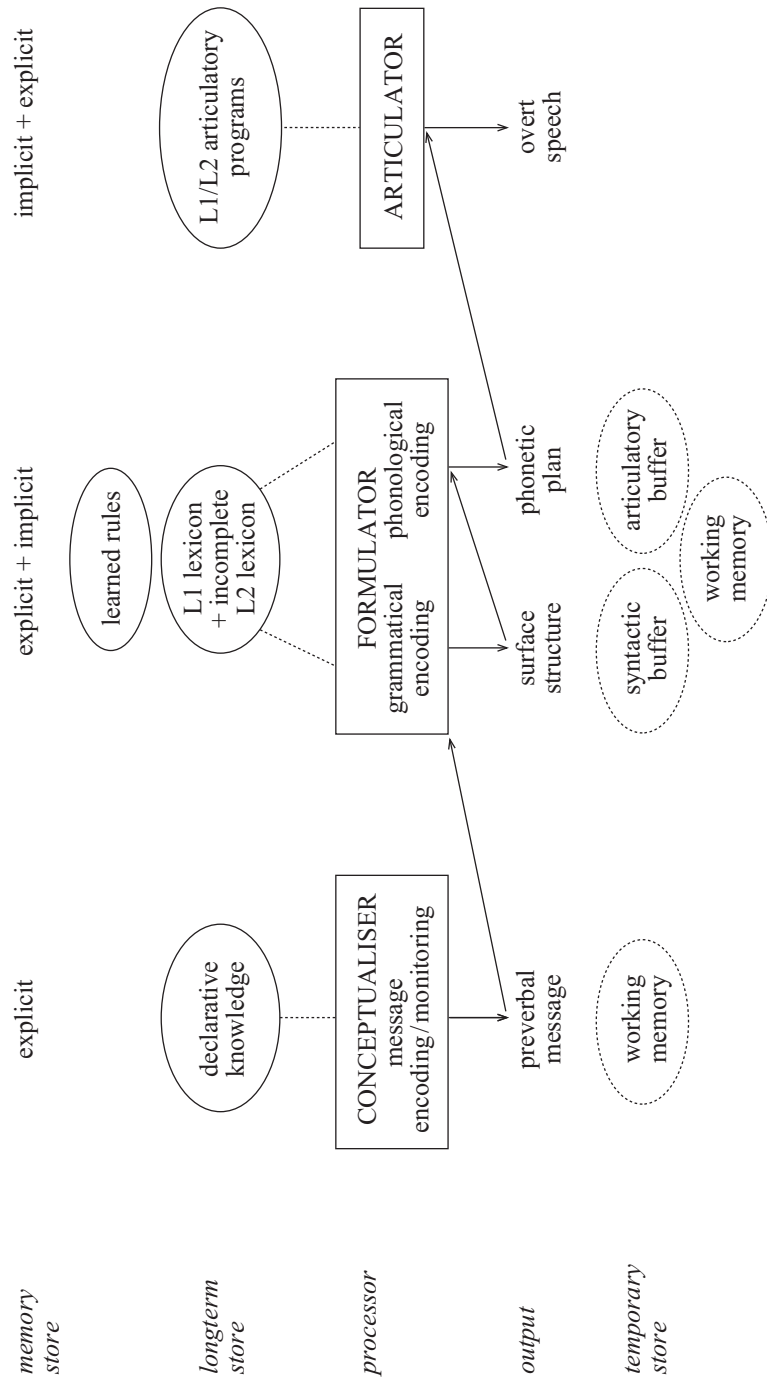


Figure 2. Nonfluent L2 Speech Production

speech production model has been adapted here (see Figure 1, adapted from Levelt's "Blueprint for the Speaker" [1989: 9]), to indicate how the different components of the speech production system responsible for each stage of processing require different sources of knowledge and different temporary stores. To resume very briefly, first the Conceptualiser produces a pre-verbal message. Secondly, the Formulator accesses the lexicon, which supplies the appropriate lemmas or lexical units, these call up syntactic procedures and are assigned to a surface structure. Next, the Formulator retrieves lexical forms for phonological encoding and constructs a phonetic plan. Finally, the Articulator outputs overt speech. What is important here is that formulating and articulating normally proceed automatically and access implicit knowledge, hence the "implicit" label of the lexicon, as opposed to Levelt's "declarative" categorisation (1989: 182). While our knowledge of word meanings as mental representations is explicit, and semantic memory is held to be explicit (Fabbro 1999: 101), the lexical-semantic system of a language is internalised as implicit linguistic competence (Paradis 1998: 427). It is normally accessed and used without conscious attention, and indeed many of the syntactic properties of lexical entries are never known explicitly (Robbins 1992: 254). In contrast, at the conceptualising stage, and in monitoring, explicit knowledge and conscious processes are used.

In the case of instructed learners, explicit knowledge – i.e. metalinguistic knowledge, such as rules, examples, translations – may be involved in all the stages of speech production. Instead of applying implicit knowledge for lexical access, lexical form building and syntactic operations, learners may use declarative knowledge acquired explicitly in the classroom or from course books. Even articulatory procedures are sometimes interrupted by the use of explicit strategies learned in order to overcome particular pronunciation difficulties. Figure 2 proposes a speech production model for non-fluent learners. The normally automatic procedures, at the level of the Formulator in particular, are replaced by qualitatively different processes if explicit knowledge is accessed; working memory – a limited capacity attentional system – may be operating at each stage, and an increase in pausing will result (Temple 2000). On the other hand, fluent learners have progressed to a level where they are no longer calling a declarative representation of syntactic procedures or of lexical forms into working memory. When these processes involve implicit knowledge and so can proceed automatically within the clause or sentence, a different balance between controlled and automatic processes is achieved – resulting in a more nativelike pause profile (see 5.1). Hence the finding of a shift to initial hesitation for this group.

### 5.5. Pedagogical implications

The first question investigated in this study was to discover whether a gain in fluency was attainable for foreign language learners within a short period of university instruction. Given the increase in speech rate by all but the three most fluent learners in the study, and given the quite substantial gains for the least fluent among them, it is apparent that it is possible for some instructed learners to improve in this aspect of fluency over a short period and without in-country experience. The lack of improvement in the most fluent learners parallels Freed's (1998) finding that even after study abroad, greater fluency (or perceived fluency) is not inevitable for such students. My study does not establish *how* learners developed fluency and a consideration of teaching methodology is beyond the scope of this article, since type of instruction was not investigated in this study. Nevertheless, the importance of input and practice must be considered. The 2-hour course unit followed by the group of learners focussed entirely on listening and speaking. Intensive listening tasks were set, all classes were held in French, and students participated in speaking activities such as exposés, discussions and debates. The students were motivated and conscious of their progress, some reporting a "break-through" in their processing of spoken input: they no longer visualised foreign words in their orthographic form, but realised they were able to access meaning directly. It is likely that meaningful practice of the two skills generally led to more fluent speakers and listeners.

Practice is, of course, the key factor in the development of procedural memory. Neurobiological research on memory informs us that long-term memories are the result of changes in neuronal structure due to experience. Furthermore, the development of procedural memory is characterised by a shift in the brain pathways used in that particular skill: "[...] neuroimaging studies appear to capture an important shift in performance from naïve to practiced conditions that reflects the operation of a procedural memory system." (Schacter et al. 2000: 637). In addition, Paradis claims that explicit knowledge cannot *become* implicit knowledge: "Not only are implicit and explicit knowledge of language subserved by different cerebral memory systems, but they have different contents, and hence one cannot become the other." (Paradis 1994: 405). The alternative brain pathway adopted in much practised performance reveals a new mode of operation where implicit knowledge is accessed automatically.

It is evident that it is through language being *used*, that phonetic, lexical, syntactic and morphological associations are gradually strengthened in long term memory and are available for automatic processing. In the case

of second language learning, it may be that automaticity is never realised to the same extent as in native speech (see Cook 1997). Experiments comparing bilinguals' performance across their two languages and also with monolinguals' performance, have found not only slower processing in the non-dominant language of bilinguals, but also slower processing in their dominant language compared with monolinguals (Dornic 1979). Recently it has been hypothesised that the crucial difference between late second language learning and first language acquisition is the use of explicit memory in L2 acquisition, involving different brain structures from those available in child language acquisition (Fabbro 1999: 100–102). This is said to explain the generally held view that acquisition of syntax is more difficult after age 6–8 years, and may never be completely nativelike (Hyltenstam and Abrahamsson 2000). However, the case of late naturalistic acquisition, in which declarative memory is less likely to be involved, is left unexplained by this account. Moreover, the results of this investigation point to a change in the balance of automatic and controlled processes, or a restructuring of the underlying mechanisms of L2 speech, and not just a speed-up of explicit memory processes.

## 6. Conclusion

While other factors undoubtedly contribute to the nonfluency of learners (such as incomplete knowledge systems, competition with first language processes, working memory capacity overload), the distinctions between implicit and explicit knowledge, and between declarative and procedural memory, are seen as especially relevant to *instructed* foreign language learning. Recalling Paradis's statement that implicit linguistic competence is acquired through interaction with speakers of the language in situational contexts, while metalinguistic knowledge is usually learned through formal instruction (Paradis 1998: 428), the question that teachers and coursebook writers as well as researchers need to address, is how can implicit linguistic competence best be achieved in the classroom? It might be instructive to discover what proportion of time in any foreign language class is spent on verbal interaction in the target language compared with acquiring and testing metalinguistic knowledge and studying written language. Interestingly, some researchers have shown that metalinguistic knowledge does not even correlate with proficiency in university language students (Elder et al. 1999).

Further research would be needed to compare, on the one hand, the effects of different types of instruction on fluency, and on the other, the fluency of



instructed learners with that of learners acquiring a second language in natural settings. In this investigation, of learners in a formal instruction setting over a relatively short period of time and without in-country experience, improvement in speech rate was achieved by all but the most fluent subjects. Greater fluency did not equate with a simple speed-up of speech, but a move to clausal planning, with less hesitant production of phrases reflecting the development of implicit knowledge and automatic procedures.

## References

- Cook, V.  
1997 The consequences of bilingualism for cognitive processing. In *Tutorials in Bilingualism. Psycholinguistic perspectives*, A. De Groot and J. Kroll (eds.), 279–299. Mahwah, N.J.: Lawrence Erlbaum.
- Dornic, S.  
1979 Information processing in bilinguals: some selected issues. *Psychological Research* 40: 329–348.
- Elder, C., J. Warren, J. Hajek, D. Manwaring and A. Davies  
1999 Metalinguistic knowledge: how important is it in studying a language at university? *ALAA* 22 (1): 81–95.
- Ellis, N. (ed.)  
1994 *Implicit and Explicit Learning of Languages*. London: Academic Press.
- Fabbro, F.  
1999 *The Neurolinguistics of Bilingualism. An Introduction*. Hove: Psychology Press.
- Freed, B.  
1998 What makes us think that students who study abroad become fluent? In *Contemporary Approaches to Second Language Acquisition in Social Context. Crosslinguistic Perspectives*, V. Regan (ed.), 89–112. Dublin: UCD Press.
- Garman, M.  
1990 *Psycholinguistics*. Cambridge: CUP.
- Garrett, M.  
1982 Production of speech: observations from normal and pathological language use. In *Normality and Pathology in Cognitive Functions*, A. Ellis (ed.), 19–76. London: Academic Press.

- Goldman-Eisler, F.  
 1968 *Psycholinguistics: Experiments in Spontaneous Speech*. London/  
 New York: Academic Press.  
 1972 Pauses, clauses, sentences. *Language and Speech* 15 (2): 103–113.
- Hyltenstam, K. and N. Abrahamsson  
 2000 Who can become native-like in a second language? All, some, or  
 none? On the maturational constraints controversy in second language  
 acquisition. *Studia Linguistica* 54 (2): 150–166.
- Koponen M. and H. Riggenbach  
 2000 Overview: varying perspectives on fluency. In *Perspectives on  
 Fluency*, H. Riggenbach (ed.), 5–24. Ann Arbor: University of  
 Michigan Press.
- Lennon P.  
 1990 Investigating fluency in EFL: a quantitative approach. *Language  
 Learning* 40 (3): 387–417.
- Levelt W. J. M.  
 1989 *Speaking. From Intention to Articulation*. Cambridge, Mass.: MIT  
 Press.
- Nation I. S. P.  
 1989 Improving speaking fluency. *System* 17 (3): 377–384.
- Paradis M.  
 1994 Neurolinguistic aspects of implicit and explicit memory: Implications  
 for bilingualism and SLA. In *Implicit and Explicit Learning of  
 Languages*, N. Ellis (ed.), 393–419. San Diego: Academic Press.  
 1998 Language and communication in multilinguals. In *Handbook of  
 Neurolinguistics*, B. Stemmer and H. Whitaker (eds.), 418–431.  
 London: Academic Press.
- Raupach M.  
 1984 Formulae in second language speech production. In *Second Language  
 Productions*, H. Dechert and M. Raupach (eds.), 115–137. Tübingen:  
 Gunter Narr.  
 1987 Procedural learning in advanced learners of a foreign language. In  
*The Advanced Language Learner*, J. Coleman, and R. Towell (eds.),  
 123–155. London: CILT.
- Robbins S.  
 1992 A neurobiological model of procedural linguistic skill acquisition.  
*Issues in Applied Linguistics* 3: 235–265.
- Schacter D., A. Wagner and R. Buckner  
 2000 Memory systems of 1999. In *The Oxford Handbook of Memory*, E.  
 Tulving (ed.), 627–643. Oxford: OUP.
- Segalowitz N.  
 2000 Automaticity and attentional skill in fluent performance. In  
*Perspectives on Fluency*, H. Riggenbach (ed.), 200–219. Ann Arbor:  
 University of Michigan Press.

- Shiffrin R. and W. Schneider  
1977      Controlled and automatic human information processing. II: Perceptual learning, automatic attending, and a general theory. *Psychological Review* 84: 127–190.
- Temple L.  
1992      Disfluency in learner speech. *Australian Review of Applied Linguistics* 15 (2): 29–44.  
2000      Second language learner speech production. *Studia Linguistica* 54 (2): 288–297.
- Towell R., R. Hawkins and N. Bazergui  
1996      The development of fluency in advanced learners of French. *Applied Linguistics* 17 (1): 84–119.
- Wiese R.  
1984      Language production in foreign and native languages: same or different? In *Second Language Productions*, H. Dechert, D. Mohle and M. Raupach (eds.), 11–17. Tübingen: Gunter Narr.



# Psycholinguistic aspects of gender acquisition in instructed GFL learning

*Barbara Menzel*

## 1. Introduction

*Child:* Die Affe nehm ich nicht mit.  
[The<sub>nom fem</sub> monkey take I not along.]  
'I am not going to take the monkey along.'

*Interviewer:* Die Affe ist bestimmt nicht richtig, weil es heißt ja nicht *die* Affe oder *das* Affe, sondern *der* Affe. Also ?  
'The<sub>nom fem</sub> monkey is certainly not correct, because it is not the<sub>nom fem</sub> monkey or the<sub>nom neu</sub> monkey, but the<sub>nom masc</sub> monkey. Well ?'

*Child:* Der Affe nehm ich nicht mit.  
[The<sub>nom masc</sub> monkey take I not along.]  
'I am not going to take the monkey along.'

*Interviewer:* Der Affe geht auch nicht.  
[The<sub>nom masc</sub> monkey works also not.]  
'The monkey doesn't work either.'

*Child:* Mhm. Was geht denn dann ?  
[Mhm. What works then ?]  
'Mhm. What will do then ?'

*Interviewer:* Mit *den*. Also, sag nochmal.  
[With the<sub>acc masc</sub>. Well, say one more time.]  
'With the. Well, try again.'

*Child:* Den Affe fährt net mit oder so.  
[The<sub>acc masc</sub> monkey drives not along or so.]  
'The monkey is not going to come along or something like that.'

*Interviewer:* Ja, dann mußt du sagen *der*. *Der* Affe fährt nicht mit, aber *den* – mit mitnehmen.

[Well, then must you say the<sub>nom masc</sub>. The<sub>nom masc</sub> monkey drives not along, but the<sub>acc masc</sub> – with take along.]

‘Well, in this case you have to say the. The monkey is not going to come along, but you have to use the with take along.’

*Child:* Warum muß jetzt immer des ich machen ?<sup>1</sup>

[Why must now always that I make ?]

‘Why do I always have to do that kind of thing ?’

This example vividly illustrates that the assignment of grammatical gender to nouns poses a challenging problem even to quite advanced learners of German as a foreign language. For a native speaker of a gender language, on the other hand, the grammatical gender of nouns obviously depicts a completely inconspicuous phenomenon in spite of its relevance for nearly every utterance, “unobtrusive to the point that native speakers may not even be aware that their language *has* a gender system” (van Berkum 1996: 2).

One of the striking difficulties the L2 learner of German has to face is the apparent arbitrariness with which the nouns seem to be distributed across the three gender classes. More than a century ago, the American writer Mark Twain has commented on this problem in his brilliantly written essay on *The Awful German Language* as follows:

A person who has not studied German can form no idea of what a perplexing language it is [...]. Every noun has a gender, and there is no sense or system in the distribution; so the gender of each noun must be learned separately and by heart. There is no other way. To do this, one has to have a memory like a memorandum book. In German, a young lady has no sex, while a turnip has. Think what overwrought reverence that shows for the turnip, and what callous disrespect for the girl.

(quoted from: Mills 1986: 12)

And although descriptive linguistics has developed a complex system of interacting rules for the assignment of grammatical gender to German nouns, based upon semantic, morphological, and phonological regularities (e.g. Wegera 1997: chap. 2), it is not clear if at all, and if yes, to what extent these regularities correspond to the psycholinguistic reality of the language learner.

Table 1. Declension paradigm of the definite article

	masculine	feminine	neuter	plural
nominative	der	die	das	die
genitive	des	der	des	der
dative	dem	der	dem	den
accusative	den	die	das	die

The problem the learner is facing is rendered even more difficult by the complexity of the declension paradigm (Tab.1). Not only does each form of the definite article reflect the fusion of information on the three morpho-syntactic dimensions gender, case, and number – the inference of a noun’s gender from the syntagmatic context is further aggravated by the high degree of syncretism within the paradigm: there are only six different forms to fill the 16 possible cells, so that no single form of the definite article defines a unique combination of gender, case, and number marking.

Since grammatical gender is not a universal feature of languages and obviously poses quite an extra burden on learners’ memories, one might thus as well ask why a language like German does not entirely do away with this feature, especially since so many other languages are functioning just fine without it. Interestingly enough, while German exhibits various symptoms of decline in the case marking system, there are no such traces whatsoever in respect to gender marking. It might thus be useful to follow Andersen’s (1984) question “What’s gender good for, anyway?” and have a closer look at the functions of grammatical gender in language processing in order to reconsider whether grammatical gender can in fact be regarded as an “unwichtiges Merkmal, das sich überall so wichtig macht” [unimportant feature that assumes an air of importance everywhere] (Werner 1975: 50).

The introductory interview passage furthermore demonstrates the elemental relevance of founded gender knowledge for the construction of the entire declension paradigm. Here, the corrective feedback of the interviewer clearly shows that the first step to finding the correct article is establishing the appropriate gender class (*der Affe* [the<sub>nom masc</sub> monkey]), followed by the identification of the article form corresponding to the appropriate case context (*der*<sub>nom</sub> resp. *den*<sub>acc</sub>) within the masculine paradigm.

Following this observation, the paper sets out to investigate the input conditions and psycholinguistic processes that influence the acquisition of grammatical gender of GFL learners in the earlier stages of language learning. Starting from the assumption that native speakers obviously make use of a complex (internal) system of gender regularities which allows them to assign gender even to previously unknown items relatively homogeneously (Köpcke and Zubin 1983: 173; Köpcke and Zubin 1984; MacWhinney 1978: 66–69; Mills 1986: 79–85), the question is to what degree GFL learners have already developed such an internal gender assignment system. The results of two gender assignment studies reveal a clear connection between correctness rates resp. latencies and type resp. token frequencies of nouns following the various gender regularities in the input. Finally, a connectionist model is constructed and trained on comparable input in order to simulate the gender acquisition process of the learners and their capability to extract gender-relevant structural patterns from the input.

## **2. Why gender?**

In many languages of this world, speakers have to supply nearly every utterance with gender marked elements whereas other – genderless – languages work just perfectly without this grammatical category. We thus have to face the question whether the language user derives substantial profit from the psycholinguistic cost connected with the acquisition of gender and the continuous processing of gender marked elements.

Several linguists (e.g. Gregor 1983; Wegener 1999; Werner 1975) estimate the functional usefulness of grammatical gender as rather scanty. Maratsos and Chalkley (1980: 188) take the existence of gender systems as evidence for unlimited cognitive resources of man, since “arbitrary gender subclass systems such as those found in Russian and German have virtually no communicative value. [...] They certainly do not reward graciously the effort to learn them by a child acquiring French, Italian, German, or Russian”. Other researchers (e.g. Bates and MacWhinney 1989; Corbett 1991; Friederici and Jacobsen 1999; van Berkum 1996) allot an important communicative function to grammatical gender in the sense that the gender marked elements significantly add to local and global text coherence.

For German, the semantic and structural functional aspects of gender marking are well documented in descriptive grammars (e.g. Eisenberg 1994: 167–176). With reference to human beings, for example, gender marked articles specify the sex of the person referred to, as with family



names in *der / die Meier* [Mr. / Ms. Meier], or with nominalized adjectives (*der / die Kranke* [the male / female sick person]) respectively participles (*der / die Abgeordnete* [the male / female member of parliament]). On structural grounds, gender marked elements facilitate discourse cohesion, e.g. through anaphoric reference: *Der Deckel<sub>masc</sub> fiel in die Kanne<sub>fem</sub> auf dem Glastablett<sub>neut</sub> und er<sub>masc</sub> / sie<sub>fem</sub> / es<sub>neut</sub> zerbrach*. [The lid dropped into the pot on the glass tray and it (???) broke.]. Although these aspects of gender function doubtlessly assist language processing, they are not completely convincing with respect to the cost-gain-relation described above.

There is, however, a third functional aspect of gender marking in language processing which has only recently attracted the interest of researchers especially in the field of psycholinguistics. A number of gender priming experiments as well as ERP (event related potentials) studies from various gender languages revealed that information on the grammatical gender of a noun is processed automatically and exerts an important – facilitative or inhibitory – influence on lexical access. The following brief description of three studies with different experimental designs will illustrate the effect of gender marking on subsequent language processing<sup>2</sup>.

For French, Grosjean et al. (1994) demonstrated how a gender marked article significantly affects lexical access of the following noun. These researchers used a gating paradigm, where spoken parts of nouns either preceded by an appropriately gender marked article or without any information on the noun's grammatical gender were presented to native speakers of French.<sup>3</sup> In the first gate, 0 milliseconds (ms) of the digitized recording of the target noun were presented to the subjects, the second gate contained the first 60 ms of the noun, and for each following gate the length was increased by 60 more ms until the end of the word was reached. The subjects were asked to write down the word they thought had been presented and to indicate how confident they were about their answer. The results of this study show that for stimuli preceded by a gender marked article, the subjects (a) needed a shorter segment of the target noun, (b) proposed less possible candidates, and (c) only proposed candidates of the same gender class (Grosjean 1994: 592–594).

Using picture-word inference stimuli, Schriefers (1993) discovered a gender distraction effect for Dutch gender marked noun phrases. In this experiment, the subjects were asked to name colored line drawings of common objects presented on a computer screen by either an article + adjective + noun combination (e.g.: *het rode huis* [the red house]) or without the definite article (*rood huis* [red house]). Together with the line drawing of the object to be named, a written word of the same semantic category but with

a different gender (e.g.: *kerk* [church]) appeared on the screen. The subjects were instructed to ignore this distracter word. Reaction time measurements revealed significantly longer latencies when target noun and distracter had a different gender. The gender distraction effect thus shows that language users automatically activate the grammatical gender of a noun even if this noun is only processed rather fugitively and is not accompanied by any gender marked elements.

For German, Schmidt (1986) conducted an experiment which is of special interest for L2 acquisition. In this study, the subjects had to perform a lexical decision task for visually presented nouns primed by (a) a gender congruent definite article, (b) a gender incongruent definite article, or (c) the neutral article substitute *xxx*. When the stimuli were presented under normal conditions, reaction times showed only inhibitory effects for incongruent gender primes relative to the neutral baseline. When Schmidt (1986) reduced the readability of the stimuli, however, congruent gender primes turned out to be especially helpful for the recognition of the target nouns. This result indicates that foreign language learners – who commonly face the problem of reduced perceptibility of the L2 input – might particularly profit from reliable gender knowledge.

Various ERP studies furthermore support the functional relevance of gender information in language processing. In these types of studies, a set of electrodes is placed on the head of the participants to measure the change in the wave forms of brain activity while ungrammatical utterances are presented to the subjects. When participants had to process gender incongruent article + noun combinations, Hagoort and Brown (1999) found a clear P600 / SPS (syntactic positive shift) effect which is also observable with other major syntactic violations.

As these studies illustrate, grammatical gender plays an important role in word recognition and production processes. L2 learners might particularly profit from gender knowledge in situations where the L2 input is not clearly perceivable. Contrary to the often expressed opinion that incorrect gender marking does not at all disturb the comprehension of an utterance, ERP studies show that incongruent gender information causes strong irritation to the recipient's processing system, especially when L2 production is rather fluent.

### **3. L2 learners of German and gender assignment**

Considering the importance of reliable gender knowledge on the one hand and the complexity of the learning task on the other, the question arises how beginning learners of German as a foreign language (GFL) deal with the problem of gender assignment. To investigate this problem empirically, natural language production was reduced in two gender assignment studies in order to allow for controlled manipulation of test items and to circumvent the problem that it is often quite difficult in German to judge whether a mistake on an article is due to erroneous gender assignment, case assignment, or even both.

#### **3.1. Study 1**

##### *Subjects*

The study was conducted with 91 university students from various faculties of two Japanese universities (A and B) who were taking German as their second foreign language for one 90-minute unit per week. The 55 participants from A-University (30 female and 25 male students) had almost finished their second year and had completed 45 units of German, whereas the 36 subjects from B-University (10 female and 26 male students) were about to finish their first year and had received 20 units of German instruction at the time of the study. All students were taught using the same textbook<sup>4</sup>, where the group from B-University had finished lesson 4 after one year, while the B-University group had proceeded until the middle of lesson 7 after two years. This textbook does not give any explicit information on existing semantic and formal gender regularities in German and no information on such regularities was given by the German teacher in class<sup>5</sup>, thus allowing to investigate what kind of gender regularities the subjects were able to deduce solely from their analysis of the input.

##### *Stimuli*

The stimuli for the study consisted of 90 items (72 real German nouns – partly yet unknown to the students – and 18 nonce words). The real nouns were chosen according to (a) the frequency of their appearance in the textbook, (b) the applicability of semantic and formal gender indicators, and (c) word-length (monosyllabic, disyllabic, multi syllabic); the nonce words were constructed considering word-length and formal gender indicators. The items were presented as written nouns in order to enhance transparency

of the formal gender indicators and to account for the prevailing teaching modality of large learner groups typical for foreign language classes at Japanese universities.

### *Procedure*

The students were tested individually in a quiet room on campus and the instructions were read aloud in Japanese by a native speaker<sup>6</sup>. The items were presented one by one and at random on the screen of a Toshiba J-3100GT laptop computer. The students' task was to indicate the gender of the nouns by naming the nominative definite article (*der, die, das*) as quickly and as accurately as possible. Since gender knowledge was defined as a function of accuracy and speed of access, latencies were measured with the help of a voicekey connected to the computer.

## 3.2. Results and discussion

### 3.2.1. Overall accuracy rates

The overall accuracy rates for the single items are displayed in Table 2. The overview shows that gender assignment even in the early stages of GFL acquisition is by no means an adhoc decision which would have yielded accuracy rates of  $\pm 33\%$  at least for the items unknown to the students. The good results for some of the unknown items (as e.g. *Danborenker* [nonce word], *Bürger* [citizen], *Reike* [nonce word]) as well as the poor accuracy rates for some of the known nouns (e.g. such as *Gemüse* [vegetables], *Bar* [bar], *Tee* [tea]) furthermore reveal that gender knowledge neither seems to depend solely on individually memorized article + noun pairs. It is rather obvious that the students have already extracted a number of formal cues relevant for gender assignment from the input.

Interestingly enough, the 1st year students generally performed better on the gender assignment task than the 2nd year students, although the latter group had received more than twice as many German lessons. Accuracy rates for all items amount 46.53% for the 1st year and to 41.25% for the 2nd year; for known items: 50.9% (1st year) vs. 43.3% (2nd year), and for unknown items: 42.2% (1st year) vs. 39.1% (2nd year). This result may well be explained by the fact that the B-University (1st year students) is one of the traditional elite institutions in Japan, whereas the ranking of A-University (2nd year students) is located in the lower middle. While other factors such as branches of studies, first foreign language (English), age of

Table 2. Accuracy rates for the individual items – Study 1

76.9%	Foto [photo]	48.4%	Bein* [leg]	34.1%	Rad** [wheel]
75.8%	Danborenker** [n.w.]	48.4%	Dilo** [nonce word]	34.1%	Rindersteak [steak]
71.4%	Frau [woman]	46.2%	Supermarkt* [supermarket]	33.0%	Pila** [n.w.]
71.4%	Stecker [plug]	46.2%	Näseperkeit** [n.w.]	31.9%	Klapptus** [n.w.]
70.3%	Mine [cartridge]	46.2%	Butter [butter]	30.8%	Film [movie]
69.2%	Kugelschreiber [ball-pen]	45.1%	Turm** [tower]	30.8%	Brötchen [breakfast roll]
69.2%	Bürger** [citizen]	44.0%	Brief* [letter]	30.8%	Käse [cheese]
68.1%	Batterie [battery]	44.0%	Bitzefisch** [n.w.]	29.7%	Kuh** [cow]
67.0%	Reike** [n.w.]	44.0%	Kindergarten* [kindergarten]	29.7%	Huhn** [chicken]
67.0%	Glühbirne [bulb]	42.9%	Krap** [n.w.]	29.7%	Medikament* [medicine]
65.9%	Langeweile** [boredom]	42.9%	Löffel [spoon]	29.7%	Bibliothek [library]
65.9%	Reiseführer** [travel guide]	42.9%	Schnupfen* [cold]	28.6%	Pläckchen** [n.w.]
64.8%	Grippe* [flu]	42.9%	Lett** [n.w.]	28.6%	Jubiläum** [anniversary]
63.7%	Herr [gentleman, Mr.]	41.8%	Wasserhahn [water tap]	28.6%	Sosch** [n.w.]
60.4%	Lehrer [teacher]	41.8%	Büro* [office]	27.5%	Kartoffel [potato]
60.4%	Fähre** [ferry]	40.7%	Buch [book]	26.4%	Brathähnchen [fried chicken]
59.3%	Kopf* [head]	40.7%	Brot [bread]	26.4%	Bank [bench, bank]
59.3%	Eis [ice]	39.6%	Gabel [fork]	25.3%	Tee [tea]
59.3%	Wote** [n.w.]	38.5%	Zahl [number]	25.3%	Faltonbar** [n.w.]
58.2%	Nänter** [n.w.]	37.4%	Waschbecken [sink]	25.3%	Bar [bar]
57.1%	Freundin* [female friend]	37.4%	Hand* [hand]	25.3%	Besserung* [improvement]
57.1%	Vater* [father]	36.3%	Wasser [water]	24.2%	Information** [information]
56.0%	Spaziergang [walk]	36.3%	Kaktus** [cactus]	24.2%	Lebensmittel** [foodstuffs]
56.0%	Haltestelle* [bus stop]	36.3%	Tür* [door]	24.2%	Griehlung** [n.w.]
54.9%	Topf [pot]	36.3%	Mosat** [n.w.]	23.1%	Skifahren* [skiing]
52.7%	Donnerstag [Thursday]	36.3%	Tinkerbrot** [n.w.]	22.0%	Spiel* [game]
50.5%	Ärztin [female doctor]	35.2%	Essen [food]	22.0%	Gemüse [vegetables]
50.5%	Campingurlaub** [camping trip]	34.1%	Weg** [way]	20.9%	Technik** [technique]
50.5%	Motto** [motto]	34.1%	Gesicht* [face]	20.9%	Schenpalarum** [n.w.]
49.5%	Kind [child]	34.1%	Zentrum** [center]	16.5%	Wahl** [choice]

Notes: \* items unknown to the first-year-students (n = 36); \*\* items unknown to all subjects (n = 91).

subjects, status of German as second foreign language, German teacher, and classroom materials are comparable for both groups, students need to achieve much higher scores in the central university entrance exam to be accepted at B-University. It is thus quite possible that the students from B-University (1st year) are equipped with more effective mechanisms for cognitive processing of information relevant for academic learning which also enable them to extract gender indicating cues from the nominal input more quickly.<sup>7</sup>

### 3.2.2. *Types of gender regularities (and exceptions)*

In order to assess the types of gender regularities the students have already established in their L2 grammar, the 90 nouns were subsumed under the following gender indicators (cf. Hoberg 1999; Meinert 1989; Wegera 1997). Since more than one gender regularity can be applied to some of the nouns, these items will be assigned to all suitable gender regularities. *Der Herr* (gentleman, Mr.) e.g. is classed with semantic regularities as well as with monosyllables.

- semantic regularities (Sem. reg.), such as *der Lehrer* [male teacher], *die Ärztin* [female doctor]
- formal regularities (Formal reg.)<sup>8</sup>, such as *der Stecker* [plug], *die Glühbirne* [bulb], *das Foto* [photo]
- monosyllabic nouns (Monosyll.), such as *der Kopf* [head]<sup>9</sup>
- nominalized verb infinitives (Nomin. verbs), such as *das Skifahren* [skiing]
- analogies (Analogies), such as *der Donnerstag* [Thursday], *die Faltonbar* [nonce word], *das Lett* [nonce word] [*das Bett* [bed]]<sup>10</sup>
- exceptions to formal regularities<sup>11</sup> (Excep. formal), such as *der Käse* [cheese], *die Gabel* [fork], *das Wasser* [water]
- exceptions to the tendency of monosyllabic nouns to take masculine gender (Excep. monosyll), such as *die Bank* [bench, bank], *das Brot* [bread]
- nouns without specific regularities concerning their gender assignment (No specific reg.), such as *der Campingurlaub* [camping trip], *die Bibliothek* [library], *das Rindersteak* [steak]

Figure 1 depicts the percentage of correct assignments for these types of gender regularities. The best results are obtained for items with semantic gender indication (60.3%) followed by nouns with formal gender indicators (48.4%) and monosyllables (42.9%). Nominalized verb infinitives (29.1%) and nouns with gender assignments in conflict with formal gender regularities (31.6%), in contrast, yield the poorest results. Subsequent t-tests show that the results for the semantic ( $t(90) = 8.1$ ;  $p < 0.000$ ) and formal regularities ( $t(90) = 4.46$ ;  $p < 0.000$ ) deviate significantly from the accuracy rate of 43.2% for all items over all subjects, whereas the correct gender assignments for nominalized verb infinitives ( $t(90) = -3.93$ ;  $p < 0.000$ ), analogies ( $t(90) = -2.59$ ;  $p = 0.011$ ), formal exceptions ( $t(90) = -6.17$ ;  $p < 0.000$ ), and exceptions to the monosyllables masc. tendency ( $t(90) = -4.4$ ;  $p < 0.000$ ) lead to results significantly below the average score<sup>12</sup>.

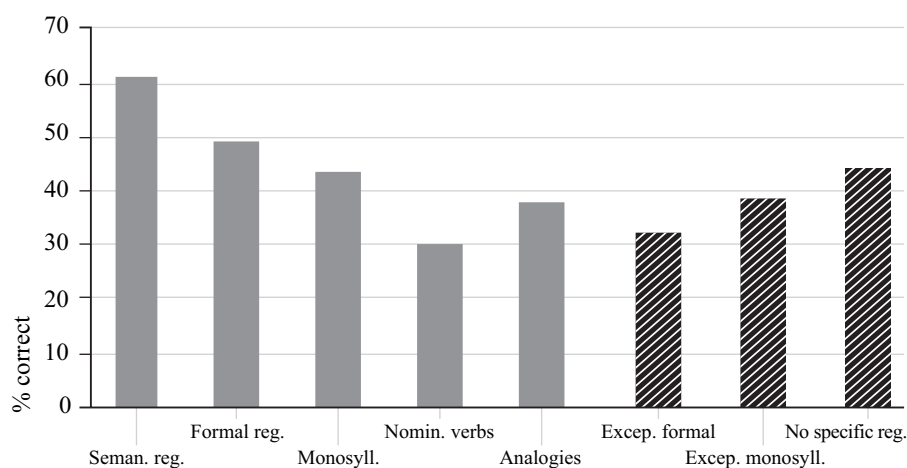


Figure 1. Types of gender regularities (and exceptions) – Study 1

These results illustrate that the subjects are obviously capable of applying semantic and formal gender regularities extracted from previous input. The top ranking of the results for nouns with semantic gender indication supports Mills (1986) hypothesis, according to which the principle of natural gender plays a more important role than could be assumed considering its relatively limited range of application in the nominal lexicon of speakers. The rather poor results for the exceptions to formal gender regularities (note that all these nouns were known to the subjects) furthermore reveal that gender assignment based upon a learner-internal system of regularities

strongly competes with memorized forms even in the early stages of language acquisition.

### 3.2.3. Formal gender regularities

We will now take a closer look at the type of formal gender regularities. Following the proposition of Altmann and Raettig (1973) and Strong (1976), the formal regularities are not further differentiated into derivational-morphological and phonological gender assignment principles, but are rather classified according to the graphematic regularities compiled by Meinert (1989: 59):

-a	fem	e.g.: <i>die Pila</i> [nonce word]
-at	neu	e.g.: <i>das Mosat</i> [nonce word]
-chen	neu	e.g.: <i>das Brötchen</i> [breakfast roll]
-e	fem	e.g.: <i>die Langeweile</i> [boredom]
-el	masc	e.g.: <i>der Löffel</i> [spoon]
-en	masc	e.g.: <i>der Schnupfen</i> [cold]
-er	masc	e.g.: <i>der Nänter</i> [nonce word]
Ge-	neu	e.g.: <i>das Gesicht</i> [face]
-ik	fem	e.g.: <i>die Technik</i> [technique]
-in	fem	e.g.: <i>die Freundin</i> [female friend]
-keit	fem	e.g.: <i>die Näseperkeit</i> [nonce word]
-ment	neu	e.g.: <i>das Medikament</i> [medicine]
-o	neu	e.g.: <i>das Büro</i> [office]
-tion	fem	e.g.: <i>die Information</i> [information]
-um	neu	e.g.: <i>das Zentrum</i> [center]
-ung	fem	e.g.: <i>die Besserung</i> [improvement]
-us	masc	e.g.: <i>der Kaktus</i> [cactus]
Ex-e	exception to -e	fem e.g.: <i>das Gemüse</i> [vegetables]
Ex-el	exception to -el	masc e.g.: <i>die Gabel</i> [fork]
Ex-en	exception to -en	masc e.g.: <i>das Waschbecken</i> [sink]
Ex-er	exception to -er	masc e.g.: <i>die Butter</i> [butter]

Figure 2 reveals major differences for the extent to which the students have established internal associations between noun endings (respectively beginning letters in the case of *Ge-*) and a certain gender class. The graph shows that these beginning students of GFL have already quite well established the word final gender indicators *-er* (65.9%), *-e* (64.3%), *-o* (54.5%),



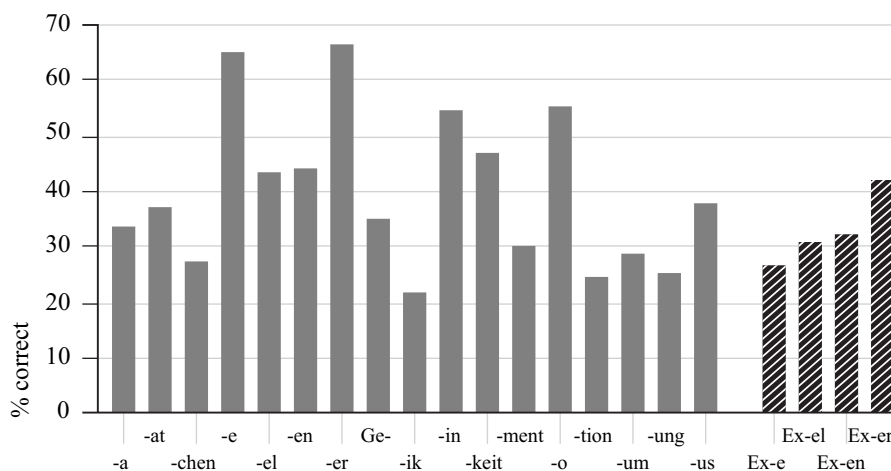


Figure 2. Formal gender regularities – Study 1

and *-in* (53.8%), whereas the highly reliable or even categorical gender indicating word endings *-ik* (20.9%), *-tion* (24.2%), *-ung* (24.7%), and *-chen* (26.8%) have not yet been acquired at all. As for the items with gender assignment in conflict to formal gender indicators, the exceptions to the *-e* fem tendency (26%) appear to be especially difficult. In order to investigate the statistical significance of the data, the result for each formal gender regularity was tested against (a) the chance value of 33.3% and (b) the accuracy rate for all items over all subjects (42.3%). T-tests yielded significant deviations for the items ending in *-e* ( $t(90) = 11.63$ ;  $p < 0.000$  for a and  $t(90) = 7.92$ ;  $p < 0.000$  for b), *-er* ( $t(90) = 12.96$ ;  $p < 0.000$  for a and  $t(90) = 9.03$ ;  $p < 0.000$  for b), *-in* ( $t(90) = 4.86$ ;  $p < 0.000$  for a and  $t(90) = 2.52$ ;  $p = 0.013$  for b), *-o* ( $t(90) = 7.05$ ;  $p < 0.000$  for a and  $t(90) = 3.74$ ;  $p < 0.000$  for b), whereas the results for *-ik* ( $t(90) = -2.9$ ;  $p = 0.05$  for a and  $t(90) = -5.21$ ;  $p < 0.000$  for b), *-tion* ( $t(90) = -2.02$ ;  $p = 0.046$  for a and  $t(90) = -4.22$ ;  $p < 0.000$  for b), *-ung* ( $t(90) = -2.65$ ;  $p = 0.012$  for a and  $t(90) = -5.52$ ;  $p < 0.000$  for b), and for *Ex-e* ( $t(90) = -2.48$ ;  $p = 0.015$  for a and  $t(90) = -5.85$ ;  $p < 0.000$  for b) are significantly lower than both test values<sup>13</sup>.

These figures show that existing formal gender indicating regularities on German nouns have been established by the students with varying success. The good results for *-e* point to a close relation between input frequency and establishment of gender regularities, as *-e* is one of the most frequent – albeit not entirely reliable – gender indicators in German<sup>14</sup>. The items ending in *-er* and *-in* might also have profited from the fact that quite a number of

these nouns refer to professions with male-female gender differentiation (e.g. *der Lehrer* – *die Lehrerin* [male resp. female teacher]) and are thus associated with the natural gender principle. In the case of *-o*, one has to take the perceptual salience of the final full vowel into account which is untypical for German and could very well lead to an early establishment of this gender regularity. The extraordinary poor results for *-chen*, *-tion*, and *-ung* are surprising, considering the fact that these suffixes categorically indicate the gender of the respective noun, and will lead to a closer look at the input frequencies of nouns with respective formal gender indicators the subjects had been exposed to.

#### 4. The role of the input

##### 4.1. Cue-category mappings

The results for the nouns with formal gender associations strongly indicate that the reliability of these regularities is not the decisive factor in the development of a learner-internal system for gender assignment. This finding is in concurrence with the predictions of the cue-category-mapping concept developed by McDonald (1989) within the framework of the *Competition Model* (Bates and MacWhinney 1987, 1989; MacWhinney 1987, 1989). One of the main ideas of this concept states that cognitive information processing implies a continuous succession of decision finding procedures with concurring candidates for each categorization task. The speaker, respectively language learner, has to evaluate each of those candidates based upon its adequateness for the linguistic decision in question. Language acquisition could thus be defined as a permanently increasing adaption to the perceived working mechanisms of the various linguistic cues in the target-appropriate application of the linguistic material. With regard to gender acquisition, the competition is one between various gender marked allomorphs for each single noun, where one of these allomorphs has to dominate the others excessively and long enough in order to ‘win out’ over the other options. The candidates are supported by cues, such as semantic and formal gender indicators, which can either add up (*der Bruder* [brother]: + sem. cue + formal cue) or compete against each other (*die Schwester* [sister]: + sem. cue – formal cue).

McDonald (1989: 375–377) developed six properties of cue-category-mappings which influence the order of cue acquisition and the strength of cue usage in language learning:

- *Decision frequency*: Cues for a frequently processed linguistic category (such as gender indicators in German) have a high general level of activation.
- *Detectability*: Salient cues should be learned before non-salient ones.
- *Availability*: The extent to which a certain cue is available in the input.
- *Reliability*: The degree to which a cue correctly indicates the classification on the cases it is present.
- *Validity*: The product of availability and reliability.
- *Conflict validity*: The extent to which a cue helps a candidate to ‘win out’ over others in case of conflicting cues.

According to McDonald’s (1989) cue-category-mapping concept, these six properties define the order in which the individual cues are being acquired, as well as their power to influence linguistic decision in cases where competent speakers are confronted with a new categorization task like gender assignment to previously unknown nouns such as loans or nonce words. On the basis of her analysis of various empirical studies on the acquisition of grammatical gender in German, McDonald (1989: 392) concludes that overall cue validity determines the acquisitional order of the different indicators to gender. Whereas cue availability constitutes the most important property in the early stages of learning, later on cue reliability gains more and more significance. Only in the late stages of acquisition, the conflict validity of a certain gender indicator becomes the decisive factor for the fine tuning of the learner internal gender assignment system.

#### 4.2. Input frequencies of gender indicators and the acquisition of gender

In order to investigate in what way the described cue-category-mapping properties account for the acquisition of gender by adult GFL learners, the results of gender assignment Study 1 (3.2.) will now be compared to the subjects’ input of nouns with the different gender indicators. Fortunately, the highly controllable nature of the acquisition process of the students who participated in the present study<sup>15</sup> allows for quite an accurate estimation of the L2 total input<sup>16</sup>. The nouns in the textbook up to the chapter which had been covered by the 1st year students at the time of the gender assignment study were thus registered according to their actual type and token frequencies and subsumed under the same gender assignment types and formal regularities described above<sup>17</sup>.

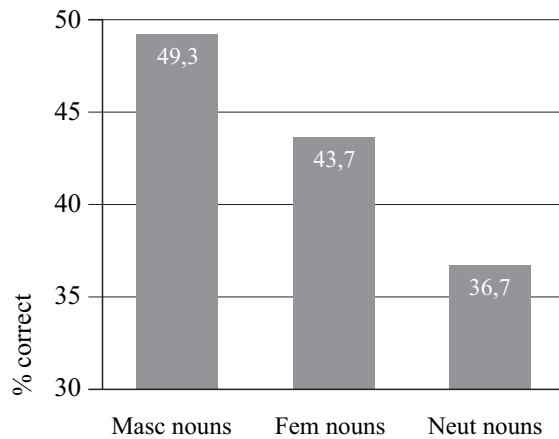


Figure 3. Correct assignments for the three gender classes – Study 1

Figure 3 depicts the overall accuracy rates for the three gender classes over all items and all subjects. Gender assignment was most accurate for the masculine items (49.3%), followed by the feminine nouns (43.7%), and the neuters (36.7%)<sup>18</sup>. T-tests show that these differences are statistically significant (masc vs. fem:  $t(90) = 3.22$ ;  $p = 0.002$ ; masc vs. neut:  $t(90) = 6.1$ ;  $p < 0.000$ ; and fem vs. neut:  $t(90) = 3.92$ ;  $p < 0.001$ ). The superiority of masculine items is somewhat inconsistent with the results of Mills (1986: 89) who found that children acquiring German as their first language perform best on feminine gender assignments. The poor results for the neuter items, however, are paralleled by an investigation with adult L1 speakers of German, where Jacobsen (1999: 507) measured significantly longer latencies for neuters compared to the other two genders in an object naming task following gender marked article primes.

The question thus arises whether the three gender classes also substantially differ in the subjects' input. Table 3 gives an overview over the type and token frequencies of nouns belonging to the three gender classes in the input to the students up to the time of the study. The masculine items which were classified most accurately in the study also have the highest type and token frequencies followed by feminine nouns, while neuters with their rather poor gender assignment were least frequent in types and tokens. The table thus clearly demonstrates how the results for the three gender classes directly correspond to the overall type and token frequencies of masculine, feminine and neuter nouns in the input.

Table 3. Distribution of the three genders in the input and assignment accuracies – Study 1

	Masculine nouns		Feminine nouns		Neuter nouns	
	Total	Ratio*	Total	Ratio*	Total	Ratio*
Types	210	1	183	0.87	161	0.77
Tokens	696	1	641	0.92	536	0.77
Correct gender assignment	49.3%	1	43.7%	0.89	36.7%	0.75

Note: \* This figure relates to the ratio of masculine : feminine : neuter nouns.

A comparison of assignment latencies, however, discloses a different picture (Fig. 4)<sup>19</sup>. The average gender assignment to feminine nouns (2193 ms) took the subjects significantly less time than to masculine (2338 ms) and to neuter (2273 ms) nouns (fem vs. masc:  $t(90) = 5.79$ ;  $p < 0.000$ , fem vs. neu:  $t(90) = -3.45$ ;  $p = 0.001$ , and masc vs. neu:  $t(90) = 2.92$ ;  $p = 0.004$ ).

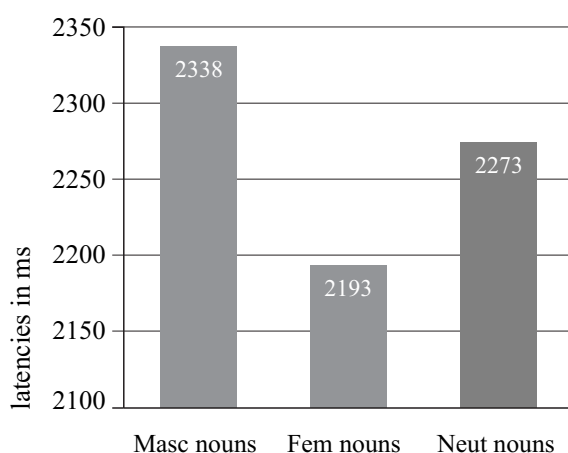


Figure 4. Latencies for the three gender classes – Study 1

A closer inspection of the distribution of the input nouns over the various formal and semantic gender regularities assists the interpretation of these findings. Table 4 considers only those nouns in the input that correspond to one (or more) of the gender regularities described above. A comparison of

the number of different gender indicators for each of the three gender classes shows that there are more gender regularities for feminine (7) than for neuter (5) and for masculine (4) nouns. At the same time, the number of types as well as tokens in the students' input which form exceptions to these gender regularities is largest for masculine nouns. We are thus confronted with the highest rule-exception-ratio for masculine nouns, second highest for neuter and lowest for feminine nouns, a difference which is clearly reflected in the assignment latencies for the three gender classes.

Table 4. Nouns corresponding to gender regularities vs. exceptions in the input

	Regular gender assignments			Exceptions		Rule-exception-ratio	
	No. of reg.	Types	Tokens	Types	Tokens	Types	Tokens
Masc nouns	4	57	180	47	149	1.21	1.21
Fem nouns	7	132	395	15	65	8.8	6.08
Neut nouns	5	20	77	8	33	2.5	2.33

This overview of overall assignment of nouns to the three gender classes seems to indicate that – at least in the early stages of GFL acquisition – the total number of nouns in the input belonging to the individual gender classes has a significant influence on assignment accuracy, whereas the latencies seem to depend more on the relation between the number of nouns that follow gender regularities compared to the number of exceptions for the respective gender class.

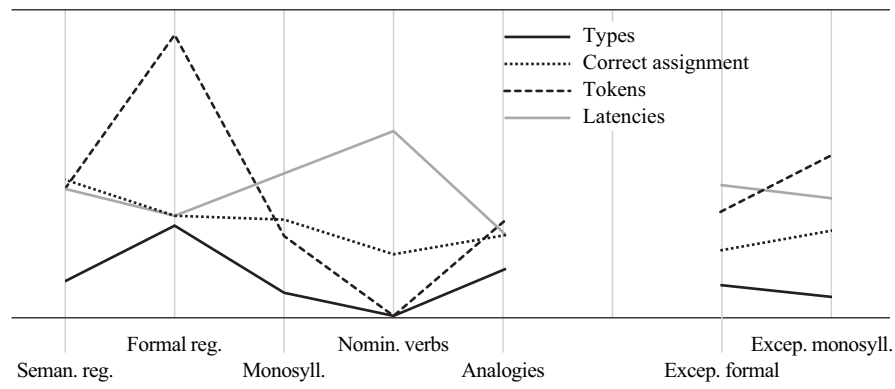


Figure 5. Influence of input types and tokens on gender assignment accuracy and latencies – Study 1

Returning to McDonald's (1989) predictions about the development of cue-category mappings in the acquisition process, we will now investigate the influence of type and token input frequencies for nouns conveying individual gender cues on the accuracy and reaction times for the corresponding items in the study. Figure 5 superposes the values for the type and token frequencies of the nouns in the input as well as for assignment accuracy and latencies for the types of gender regularities described above<sup>20</sup>. The graph depicts that the token frequency of nouns in the input intersects counter-symmetrically with the respective assignment latencies. A high token frequency is thus reflected in shorter reaction times, while types of gender regularities with relatively few tokens yield longer latencies. The values for the number of noun types following a certain gender regularity in the input, however, run basically parallel to the accuracy of gender assignment, indicating that high type frequencies for a certain gender regularity are paralleled by good gender assignment results to the nouns carrying the respective gender cue. The only exception to this pattern are the semantically motivated gender assignments. Here, the conceptual salience of the association between a semantic feature of the referent and the grammatical gender of the corresponding noun obviously strengthens these gender cues in spite of a relatively low type frequency in the input.

Figures 6a and 6b illustrate that this general tendency of concurrence between token frequency and assignment latencies on the one hand and between type frequency and accuracy on the other hand also holds for the various formal gender regularities<sup>21</sup>. Although there are some minor deviations from the described pattern (e.g. for the nouns ending in *-keit*, *-tion*, *-us*, or *-ung*)<sup>22</sup>, the data in general support McDonald's (1989: 392–393) claim

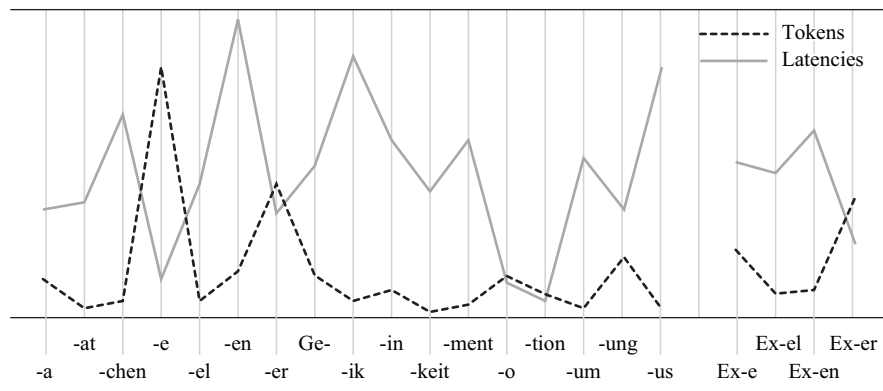


Figure 6a. Influence of input tokens on assignment latencies – Study 1

that cue availability (i.e. frequency) is the most decisive factor for the development of cue validity in the early stages of language acquisition for an instructed L2 learning setting as well.

The results of the described study on gender assignment by adult students of German in an instructed learning environment thus show that even beginning GFL learners have already extracted some of the most valid gender regularities from the input and are applying this knowledge to the gender categorization of unknown items. In the early acquisitional stages, input frequency and availability of the individual gender indicators are obviously more important than their absolute reliability for establishing cue validity. The poor results for nouns forming exceptions to the most valid gender regularities for these learners (the association of word final *-e* with feminine and of *-er* with masculine gender) indicate that at this learning stage, gender cue strength is not yet influenced by the concept of conflict validity, as these gender cues even dominate deviant individual nouns with high input frequencies (see Fig. 2). A comparison of type and token frequencies for input nouns conveying certain gender regularities with the accuracy rate and the assignment latencies in the study furthermore demonstrates the tendency of high token frequencies to correlate in particular with shorter latencies, while high type frequencies are more closely related to gender assignment accuracy. These findings correspond to the prediction from Bybee's (1988, 1995) Network Model, according to which the type rather than the token frequency of a morphological pattern is the decisive factor for its productivity.

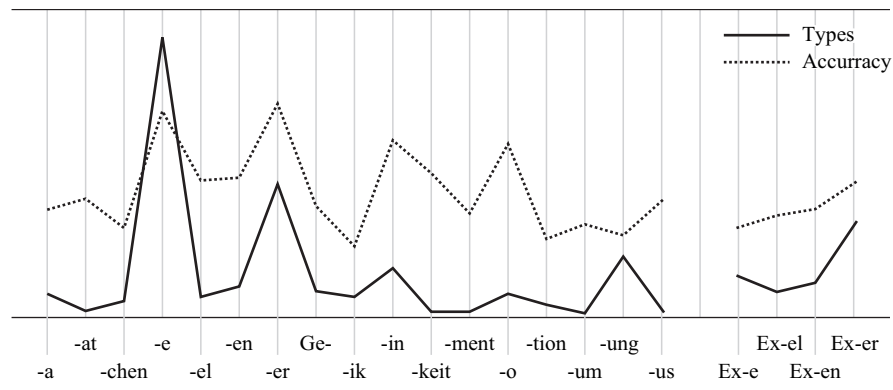


Figure 6b. Influence of input types on assignment accuracy – Study 1



#### 4.3. Resolution of semantic vs. formal gender cue conflicts

Although the development of cue conflict validity marks the last stage in the acquisition of cue-category mappings in McDonald's (1989) concept, GFL learners nevertheless do have to deal with conflicting gender indicators from the very beginning of their language learning process on. Here, conflicts between semantic and formal regularities (e.g. *das Mädchen* [girl]: + formal cue, – semantic cue; *der Junge* [boy]: + semantic cue, – formal cue) are of special relevance for beginning learners. Depending on the language to be learned, the resolution of semantic vs formal gender cue conflicts will be biased in favor of the type of gender regularity predominant in that language. In a series of experiments, Karmiloff-Smith (1979) demonstrated an unequivocal dominance of phonological over semantic regularities for French. For German, however, the picture is not that clear. For native speakers at least, Köpcke (1982: 111) places the semantic gender regularities at the top of his gender cue hierarchy, a claim that is supported by a nonce word experiment conducted by Lang (1976). In this study adult speakers of German first assigned gender to nonce words on the basis of formal gender indicators, but later changed some of these gender assignment when the nonce words were given meanings associated with a certain gender, such as flower names (fem), alcoholic beverages (masc) or names of dances (masc). For children learning German as their L1, however, Mills (1985, 1986) found that both semantic and formal gender regularities are acquired simultaneously and that no gender cue type dominates over the other, a claim supported for children learning German as their L2 in an error analysis study by Pfaff (1992).

##### 4.3.1. Study 2

In order to investigate how adult learners of GFL tackle the problem of conflicting gender cues, a study was conducted in which the stimulus items represented male and female persons, while the validity of the formal gender indicators on the nouns was varied systematically.

##### *Subjects*

The subjects were the same student groups who participated in Study 1 (2.1).

*Stimuli*

16 nouns that name persons or professions (8 male, 8 female) were selected and supplemented by drawings representing these persons. The formal gender regularities on the nouns either reinforce semantic gender indication (*der Kellner* [waiter]) or created a conflict between semantic and formal gender cues (*der Kunde* [customer]). In order to control the influence of memorized article + noun combinations, half of the stimuli consisted of items known to all 91 subjects and the other half were constructed nonce words. Gender indicators were distributed over the stimulus items as follows: 4 fem nouns conveyed convergent gender cues (-e and -in, 2 items each), 4 fem nouns conveyed conflicting gender cues (-er and -us, 2 items each), 4 masc nouns conveyed convergent gender cues (-er and -or, 2 items each), and 4 masc items conveyed conflicting gender cues (-e and -is, 2 items each).

*Procedure*

The influence of the natural gender of the referent on the assignment of grammatical gender to the respective noun can best be ascertained by a direct comparison of gender assignment to (a) solely the nouns and (b) the nouns together with the visual depiction of the referent. In order to avoid presenting each item twice to the subjects, a counter-balanced design employing two test sheets was chosen. Test sheet 1 included the items 1–8 as written nouns only, and gave nouns 9–16 together with the picture of the referent. Test sheet 2 presented items 1–8 together with the drawings and items 9–16 as written nouns only. The two test sheets were randomly issued to the students in both learning groups, and the subjects' task was to indicate the appropriate grammatical gender of the nouns by supplying the nominative definite article (*der, die, das*) on the test sheets.

*4.3.2. Results and discussion*

Figures 7 a + b show the results for the items with congruent semantic and formal gender cues. It is obvious that the already well established formal gender regularities (75% correct answers for masc items and 64% correct answers for fem items without picture) are reinforced when the formal gender cue is furthermore supported by the visual depiction of the referent's natural gender (86% correct answers for masc nouns resp. 80.5% for fem items with picture).

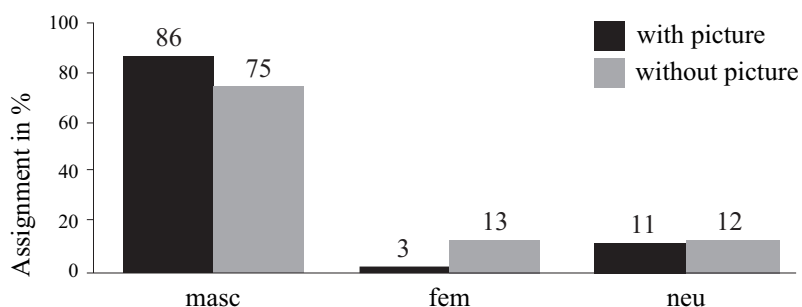


Figure 7a. Convergent gender cues; masc items – Study 2

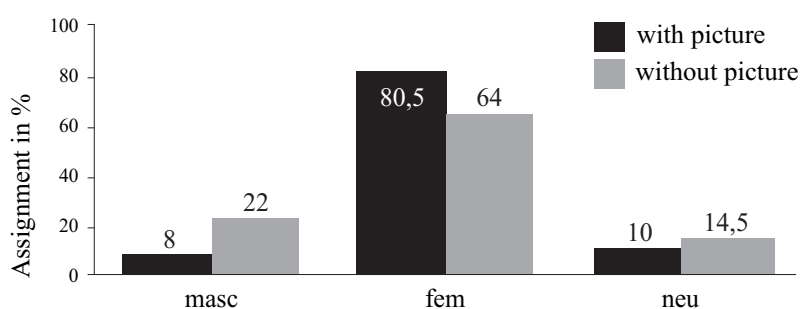
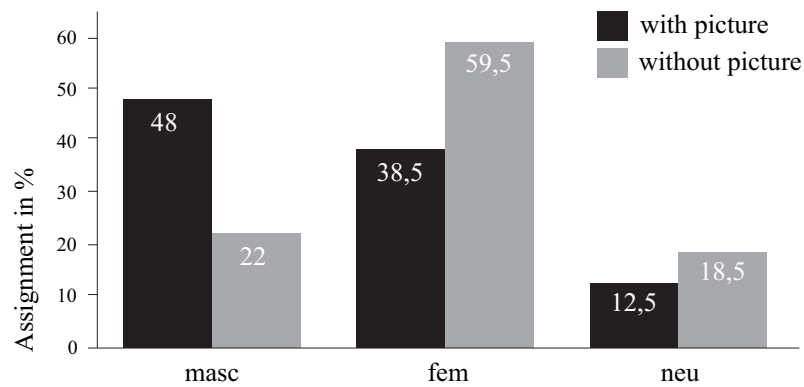


Figure 7b. Convergent gender cues; fem items – Study 2

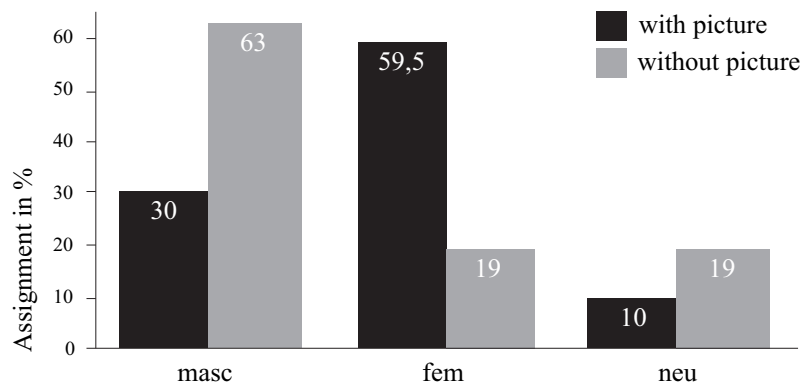
If, however, semantic and formal gender indicators compete with each other, the results turn out quite differently. As shown in Figure 8a, subjects are quite strongly oriented towards formal gender cues (59.5% of answers classified masc nouns as feminines), when known (sic!) items were presented without visual depiction of the referent. And even when a picture clearly indicates the natural gender of the referent as male, 38.5% of answers still assigned fem gender to those items.

A similar picture was obtained for the feminine items with conflicting gender cues (Figure 8b). The somewhat less dominant influence of the formal indicator of word final *-er* for masc gender when the pictures clearly identified the referents as female persons (30%) can be attributed to the fact that *-e* fem obviously has been established more strongly by the subjects than *-er* masc, following cue input availability defined by the cue-category-mapping concept (McDonald 1989; cf. chap. 4.2).

Based on these results, the resolution of semantic vs. formal gender cue conflicts by beginning GFL learners can thus be recapitulated as follows:



*Figure 8a.* Conflicting gender cues; masc items – Study 2



*Figure 8b.* Conflicting gender cues; fem items – Study 2

In case of congruent semantic and formal gender indicators, both types of cues reinforce each other. If there is a conflict between both types of gender cues, the decisive criterion for the resolution of the conflict is the strength of the formal gender regularity. Cues with high availability in the learners' input tend to dominate the semantic gender indication at this stage of GFL acquisition, whereas semantic cues dominate formal regularities with low input availability<sup>23</sup>. These observations correspond to Mills (1986: 113) findings from L1 acquisition of German: "In conclusion, an interaction between semantic and formal rules must be postulated which depends not on the categorization of the rules as semantic or formal, but rather on the relative 'clarity' of the rules in question within the gender system."

## 4.4. Formal gender cue conflicts

While the resolution of semantic vs. formal gender cue conflicts can be investigated directly by comparison, it is much more difficult to assess how students deal with inconsistent formal gender indicators. Following the acquisitional order of cue-category mappings in McDonald's (1989) approach, cue conflict validity should be the last property to be learned, so that it could be assumed that beginning learners of German have not yet come to terms with this problem at all. A closer inspection of the latencies for the items of Study 1, however, reveals that a certain degree of inconsistency of some formal gender regularities is well reflected in the time subjects need to assign gender to the items in question. Figure 9 shows that the nouns with word final *-en* (regular items as well as exceptions) yielded the longest latencies (2620 ms resp. 2382 ms)<sup>24</sup>. Note that the accuracy rates for these nouns do not deviate significantly from the average results over all items. In accordance with the assumption that the proportionally large number of exceptions to existing gender regularities for masculine nouns might be the reason for the longest latencies for this gender class (Table 4), the problem of cue conflict validity for individual formal gender indicators could have caused longer latencies for the assignment of gender to the nouns following this regularity. In the case of *-en*, the association with masculine gender is relatively reliable, as 72.1% of nouns ending in *-en* are masculine according to the list of basic vocabulary (Wegener 1999: 513). In the subjects' input, 10 types with 37 tokens followed this gender regularity. At the same time, however, there are not only some 'real' exceptions, such as *das*

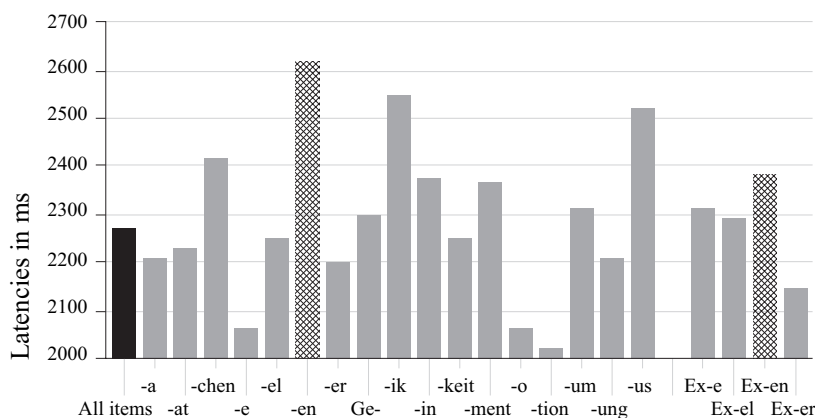


Figure 9. Latencies for formal gender regularities – Study 1

*Waschbecken* [sink], but at least two conflicting gender regularities with word final *-en*: all nominalized verbs derived from the infinitive form end in *-en* and have neuter gender, e.g. *das Skifahren* [skiing], and the diminutive form *-chen* (*das Mädchen* [girl]) also ends in *-en* and categorically assigns neuter gender. Furthermore, the plural form of most feminine nouns ending in *-e* take the plural suffix *-n*, so that students will be confronted with final *-en* on a large number of nouns classified as feminines as well.

One can thus conclude from the latency results that learners are already working on the problem of gender cue conflicts even in the earlier stages of GFL acquisition, as is reflected in significantly longer reaction times for items with inconsistent formal gender indicators. Due to the complexity of the task, however, this property of cue-category-mappings will be mastered only in the later stages of the acquisition process.

## 5. Connectionist simulation

The analysis of the gender assignment data for GFL students as well as the comparison with the input frequencies for nouns with the various gender regularities indicates that language learners obviously scan formal gender indicating properties of nouns and thus deduce gender class relevant information from the input. This makes the acquisition of grammatical gender an ideal test case for the explanatory power of connectionist models in language processing, and in recent years a number of such networks have been developed for several gender languages (e.g. Taraban et al. 1989, and MacWhinney et al. 1989, for German; Sokolik and Smith 1992, for French; Taraban and Kempe 1999, for Russian). By simulating only selected aspects of the learning process, all these models necessarily simplify the learning task considerably, but nevertheless show striking parallels with empirically ascertained phenomena of the course of language acquisition. MacWhinney et al. (1989) e.g. compare the performance of their connectionist network on the acquisition of the definite article declension paradigm in German with MacWhinney's (1978) and Mill's (1986) empirical data on German L1 acquisition and conclude that "the model provided a good match to the currently available data on the acquisition of the declension of the definite article in German" (MacWhinney et al. 1989: 274).<sup>25</sup>

After a brief introduction to the working mechanisms of connectionist models, the following chapter will present a simulation of the L2 German gender acquisition process and compare the performance of the model with the study data described above.

### 5.1. Working mechanisms of connectionist models

The term ‘connectionist model’ was coined by Feldmann (1981) as a specification for cognitive models aiming at an augmented incorporation of detailed information on the neuronal architecture of the brain into modelling. Here, the concept of ‘connection’ explicitly refers to the massive parallelism of neuronal processes in multiply entwined networks. In terms of instructed second language acquisition research, the increased interest in connectionist approaches is related to a shift in focus on the acquisition *process* rather than the acquisition *product*. Using computer technology, connectionist simulations attempt to recapitulate the various steps in the development of linguistic structures, as well as to show how learners employ their integrated cognitive architecture to deduce and analyze input patterns in a manner which enables them to build structural regularities without having to resort to symbolic rules.

It is important to note that connectionist models by no means claim to depict small-scaled *replicas* of the complexity of human cognition. In their function to serve as *models* for selected aspects of cognitive processes, they are constructed to merely simulate information processing operations of neuronal networks in the brain. In these networks of parallel distributed processing (PDP), information is manifested in the activation patterns of interconnected neurons (‘units’ in connectionist models) – there is no symbolic information in the individual units themselves. During language learning, ongoing system-internal input analysis causes permanent upgrading of the activation values within the network and thus leads to a gradual extraction of probabilistic patterns of the structural regularities of the perceived language input. While these activation patterns are inevitably rather instable in the early stages of acquisition, one should expect a progressive approximation to target structures in later learning stages. In terms of gender acquisition, this model would predict an earlier ‘settling’ of activation patterns for gender regularities with high input frequencies and few exceptions (reflected in good accuracy rates and short latencies for the assignment of gender to the respective nouns in the study data). Activation patterns for gender cues with low input frequencies or high assignment conflict potential, on the other hand, would be subjected to frequent ‘upheavals’ in the activation values and should thus lead to delayed settlement of stable patterns (reflected in low accuracy rates and longer latencies for gender assignment to the nouns in question in the study data).

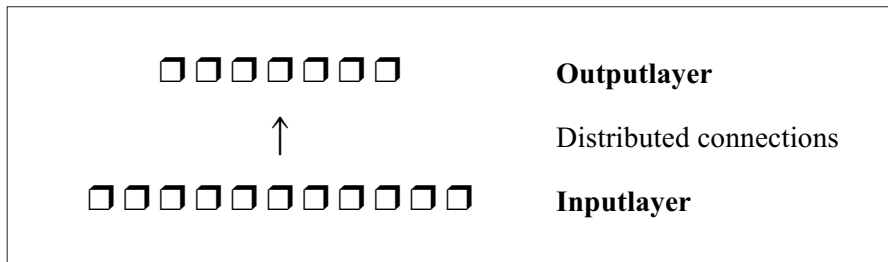


Figure 10. Two-layered perceptron

For the construction of a connectionist model we thus need information on the input, on the output, and on the type of connections between the input and the output of the system. The simplest variants of such models are two-layered perceptrons (Minsky and Papert 1969; see Figure 10). Applying a simple learning algorithm, these models are capable of learning any task with linearly separable solutions. Many complex cognitive learning tasks, such as e.g. Boolean XOR<sup>26</sup>, however, can not be represented in a linearly separated manner and can therefore not be solved by a perceptron. In order to tackle more complex tasks, the model needs to be expanded by one or more inserted layers, so-called *hidden* layers (Rumelhart and McClelland 1986), which allow for model-internal representations of the extracted patterns and thus lead to restructuring of the available solution space. Figure 11 depicts a simplified scheme of a multi-layered model with one layer of hidden units.

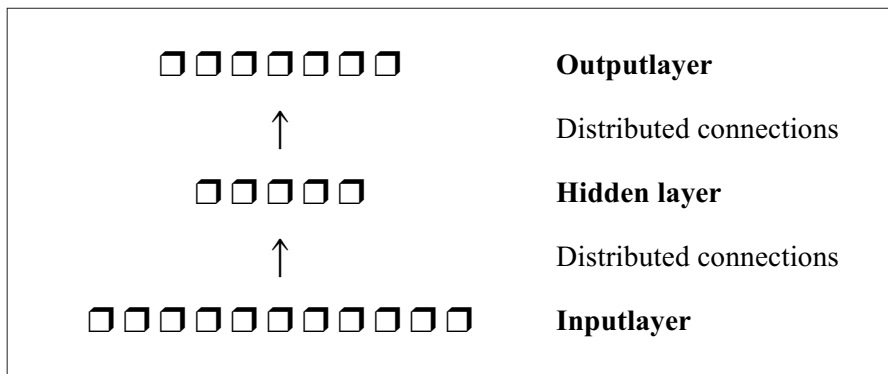


Figure 11. Multi-layered connectionist model



While learning in a network without hidden layers corresponds to the conventional statistical principle of multiple linear regression, multiple-layered networks employ distributed computational algorithms that go beyond the limitations of traditional statistical operations. In PDP model variants developed by Rumelhart et al. (1986), learning is based upon the *backpropagation* algorithm which allows for computational solutions of non-linear problems. This learning algorithm is responsible for the continuous adjustment of the activation values in the network according to upgraded input information. For each input pattern presented to the network, the distributed connections produce a certain activation pattern on the units of the output layer. This output activation is then compared to the expected output pattern (i.e. the target-like form in case of a language learning task). If the network finds a discrepancy between the network generated and the expected output patterns, it will use this mismatch to change the activation values on the connections between output units and hidden units as well as on those between hidden units and input units. In this way, the error message is ‘backpropagated’ through the network and the activation values are adjusted to ensure that the actual output will correspond better to the expected output the next time the network is presented with the same or a similar input pattern<sup>27</sup>. It is important to note that all input patterns are processed in the same way, i.e., there is no special treatment (such as the application of pre-programmed if-then rules) for regular vs. irregular forms in language learning tasks like gender assignment. Since all possible classifications have to be processed by the same network constellation, the error-signal-based adjustment of activation values has to proceed in very small steps, so that one single unexpected classification, as e.g. *der Affe* [monkey] in a gender assignment learning task, will not cause the network to classify all following items ending in *-e* as masculines. It is the layer of hidden units which enables the network to restructure the solution space in such a way that ‘deviating’ classifications can gradually be incorporated without hard-wired rules for ‘exceptional cases’.

The described multi-layered model architecture with the backpropagation learning algorithm is certainly not the only possible structure for a connectionist model. There are numerous other approaches with different network architectures and learning principles<sup>28</sup>. Previous connectionist networks (e.g. MacWhinney et al. 1989; Taraban et al. 1989; Taraban and Kempe 1999), however, have demonstrated that PDP models employing backpropagation seem to be especially well-suited for the simulation of language acquisition tasks, moreover so in respect to instructed language learning settings where learners are often supplied with direct corrective feedback.

## 5.2. Simulation of the data from Study 1

The network simulator T-LEARN was employed to construct a connectionist model of German L2 acquisition of gender assignments<sup>29</sup>. During model training, a series of German nouns was presented to the network on the input layer, and the network's learning task was to activate one of the three output units (masc, fem, or neu) according to the gender class of the respective noun. After each classification, the actual output was compared to the target output while the network backpropagated the error message through the system in case of mismatch to adjust the connection patterns accordingly.

It should be noted that the goal is not to construct a simulation that manages to learn German gender assignments in a most effective manner, but to try to simulate the learning process of the students involved in Study 1 as closely as possible. Therefore, the network's training set of input items consisted of those nouns that had appeared in the textbook up to the time of the study<sup>30</sup> according to their actual token frequencies, so that the item *der Käse* [cheese] e.g. was represented eleven times in the training set. In order to ensure correspondence between network input and learning experience of the students, the coding of the target output was determined following the gender classification of the nouns in the textbook, even though the network was thus confronted with some unsolvable problems: Some items with unstable gender (such as *der/das Bonbon* [candy], *der/das Joghurt* [yoghurt]) were double-coded, and with homonyms such as *der See* [lake] vs. *die See* [ocean], the input coding offered the network no option for disambiguation, as no semantic properties of nouns were included.

Input-coding:					
# =	0	0	0	0	0
a =	0	0	0	0	1
ä =	0	0	0	1	0
b =	0	0	0	1	1
c =	0	0	1	0	0
d =	0	0	1	0	1
...					

Figure 12. Input-coding

Following these criteria, the training set comprised 1859 items which – considering the fact the the dominant input modality for Japanese students in large foreign language classes is written language – were coded according to their orthographic form. The longest input items consist of 14 letters, and

each letter is represented by a 5-bit binary code (see Figure 12), so that the input layer of the network is made up of a 70-bit input vector. Since all units on the input layer have to be filled for each input, nouns with less than 14 letters were coded in right alignment with # as filler for the missing letters. This coding principle automatically supplies the network with the information that word endings are more crucial than word initial letters for the learning task ‘gender assignment’.<sup>31</sup> Some legitimation for the chosen procedure can be obtained from a questionnaire in which 67 out of the 91 subjects in Study 1 stated that they attend to noun endings when assigning grammatical gender (Menzel 2004: 136).

For the learning task at hand, the number of units on the output layer was fixed to be three, one for each gender. A less obvious decision, however, is the estimation of the number of units on the hidden layer(s), which has to be determined by the model constructor when working with T-LEARN<sup>32</sup> and is usually ascertained through a number of pilot simulations. As explained above, the model needs these hidden units to be able to restructure the available solution space in order to solve problems that are not linearly separable. With very few hidden units, the network might not have enough processing capacity to extract general input patterns. If the number of hidden units is very large, however, the network has much memory capacity and will learn the training set very quickly, but will not be able to generalize well and will thus perform poorly on the classification of previously unseen items. A series of pilot simulations found the described network to work best with one hidden layer of 20 units, so that the model architecture consists of 70 input units, 20 hidden units, and 3 output units (see Figure 13).

During the network’s learning phase, the complete training set of 1859 randomly ordered nouns was presented to the model five times to account

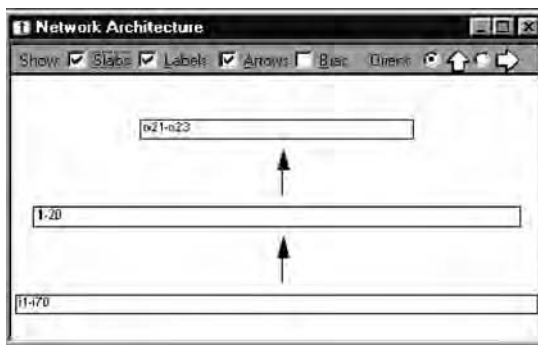


Figure 13. Network architecture

for repetition in class, homework and exam preparation of the students. One measurement of network performance after training is the RMS (Root Mean Square) error curve<sup>33</sup> which documents the average error ratio over all input patterns. Figure 14 depicts that this global network error value is about 0.4 after five runs through the training set, i.e., the model achieved an accuracy rate of 60% on all training items.<sup>34</sup>

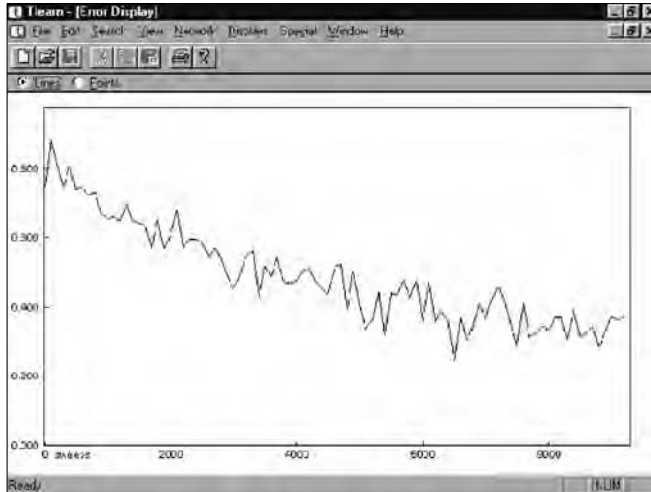


Figure 14. Error display

Of much greater interest than the RMS error curve, however, is to test how the trained network is capable of generalizing the extracted input patterns to previously unseen items. For this purpose, the configuration of the network connections between input layer, hidden layer and output layer is 'frozen' after the end of training. The network is then confronted with a test set of new nouns and has to categorize these items on the basis of the frozen connection values without any further corrective feedback.

There is a relatively good match between subjects' assignment results and network performance for most of the items. Nouns that had yielded accuracy rates above average in Study 1 (e.g. *das Foto* [photo], *die Glühbirne* [bulb], *der Kugelschreiber* [ball-pen]) achieve high activation on the appropriate output units, and many of the items that were classified exceptionally poorly in the study (e.g. *die Bank* [bench, bank], *das Spiel* [game], *die Bibliothek* [library]) show either inappropriate or generally low output unit activation in the simulation. Nevertheless, the correspondence between study results and network performance is by no means perfect, as there are

several items that yielded poor accuracy rates in the study but were properly classified by the network (e.g. *der Tee* [tea], *das Gemüse* [vegetables]).<sup>35</sup> Furthermore, there are some items for which network classification was poorer than the subjects' assignment accuracy rates (e.g. *die Ärztin* [female doctor], *die Freundin* [female friend], *der Vater* [father]). Interestingly enough many of these items are nouns with semantic gender cues, so that this result is not surprising considering that – unlike the students – the network had no information whatsoever on semantic features of the nouns.

Table 5. Gender assignment accuracy (Study 1) and network performance

Test item	Gender class	Accuracy of assignment in Study 1	Activation of network output units		
			<i>masc</i>	<i>neu</i>	<i>fem</i>
Ärztin	fem	50.5%	0.485	0	0.737
Bank	fem	26.4%	0.815	0.261	0.01
Besserung	fem	25.3%	0.449	0.024	0.291
Bibliothek	fem	29.7%	0.285	0.314	0.27
Campingurlaub	masc	50.5%	0.984	0.144	0
Eis	neu	59.3%	0.06	0.982	0
Foto	neu	76.9%	0	0.996	0
Frau	fem	71.4%	0	0.068	0.964
Freundin	fem	57.1%	0.961	0.015	0.11
Gabel	fem	39.6%	0.789	0.898	0
Gemüse	neu	22.0%	0.025	0.932	0.149
Glühbirne	fem	67.0%	0.101	0	0.958
Herr	masc	63.7%	0.996	0.076	0
Kartoffel	fem	27.5%	0.465	0.13	0.454
Kugelschreiber	masc	69.2%	0.985	0.055	0
Mine	fem	70.3%	0.022	0.067	0.669
Spiel	neu	22.0%	0.61	0.532	0.022
Stecker	masc	71.7%	0.983	0.124	0
Tee	masc	25.8%	0.602	0.01	0.019
Vater	masc	57.1%	0.643	0.611	0

T-LEARN offers a number of options to analyse the model-internal representations of the co-occurrence patterns the network built up during training<sup>36</sup>. First of all it is important to examine the strength of the individual connections between the units of the different layers (input, hidden, and output) to ensure that categorization is not due to extremely high activation on some units only, which would indicate that the number of hidden units is too large to force the network to extract patterns that accommodate all possible inputs.

Figure 15 depicts part of the weight file after the five training runs and clearly shows that none of the connections has assumed an over-proportionally high value, so that the activation of the output units is in fact due to the ensemble of the complete distributive set of connections.

An other option is to visualize the pattern representation on the layer of hidden units for each item where white units indicate strong activation and grey units weak activation. Figure 16 illustrates the activation pattern for the item *die Glühbirne* [bulb] on the input layer (bottom row), on the hidden layer (middle row), and on the output layer where the unit on the right represents the categorization for feminines.

```

Tlearn - [gender-9295]
File Edit Search View Network Disp
NETWORK CONFIGURED BY TLEARN
# weights after 9295 sweeps
# WEIGHTS
# TO NODE 1
+0.2537763119
-0.6365693331
0.7732062163
-0.6555174494
-0.0342406452
-0.8359666753
0.1337285726
-0.8509421849
0.4912436604
-0.4379656899
-0.8464483397
-0.8448867202
0.7137049585
0.0274810418
-0.6592572331
-0.3305580902
-0.6367813468
-0.5352853251
0.3621477783
0.0258183628
0.5536712794
-0.8001190424
-0.4351645732
-0.3794060946
-0.7450715303
-0.0335147991
Ready

```

Figure 15. Weight file

represents the categorization for feminines.

By comparing this visualization of node activity patterns for items with the same gender cue, one could now investigate whether the network has extracted a common activation pattern e.g. for all feminine nouns ending in *-e* and contrast this pattern with the one for ‘exceptions’, i.e. masculine nouns ending in *-e* such as *der Käse* [cheese].

A cluster analysis of the activation values on the hidden units for each item (Figure 17) finally illustrates how the network has restructured the multi-dimensional solution space for the set of test items<sup>37</sup>. The cluster display shows that the network first separates the items classified as neuters from those classified as masculines and feminines, and in a second step differentiates between masculine and feminine nouns. Subsequent branching then divides the nouns in each gender class into subgroups of related items

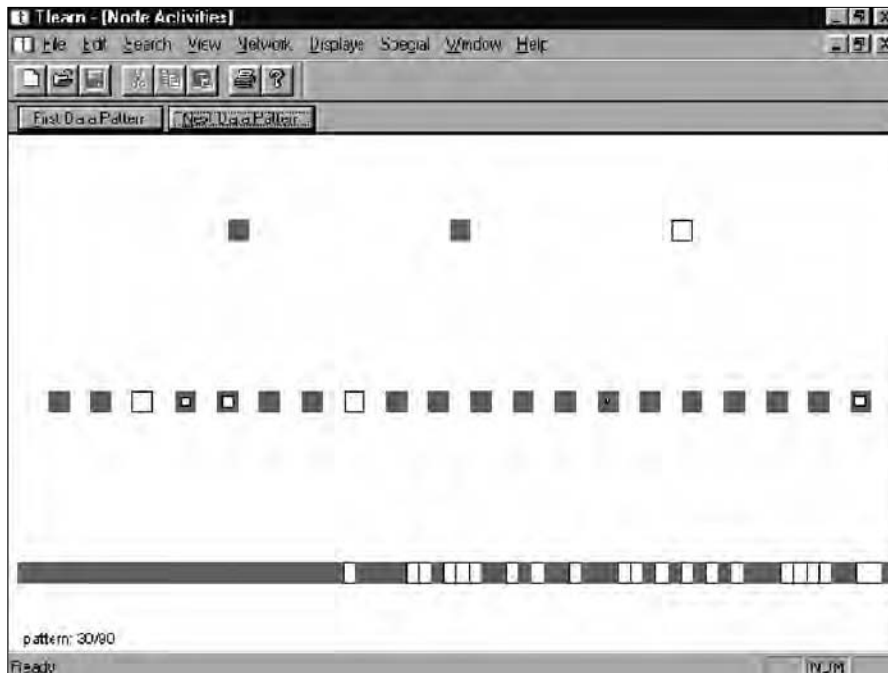


Figure 16. Node activities (example: *Glühbirne* [bulb])

according to the structural pattern representations that were built up by the network during training.

The connectionist network presented above has thus achieved a fairly good simulation of the results of the study data, although the only information available for item classification was the orthographic structure of the nouns. It is important to note that the network's achieved 'knowledge' of German gender assignment does not rely upon pre-programmed if-then rules, but is based upon a net of distributed connections between a large number of units on the different layers of the model. Unlike in the connectionist simulation, *in vivo* such networks certainly do not operate in isolation from other language processing operations: Semantic aspects of the nouns as well as already established connections within the declension paradigm of the definite article are automatically activated by the language learner and contribute to the task of gender assignment. On the other hand, machine learning is not afflicted by external, affective and other individual factors that influence human learning, such as attendance span, fatigue, and motivation. In spite of these aspects responsible for some deviation between



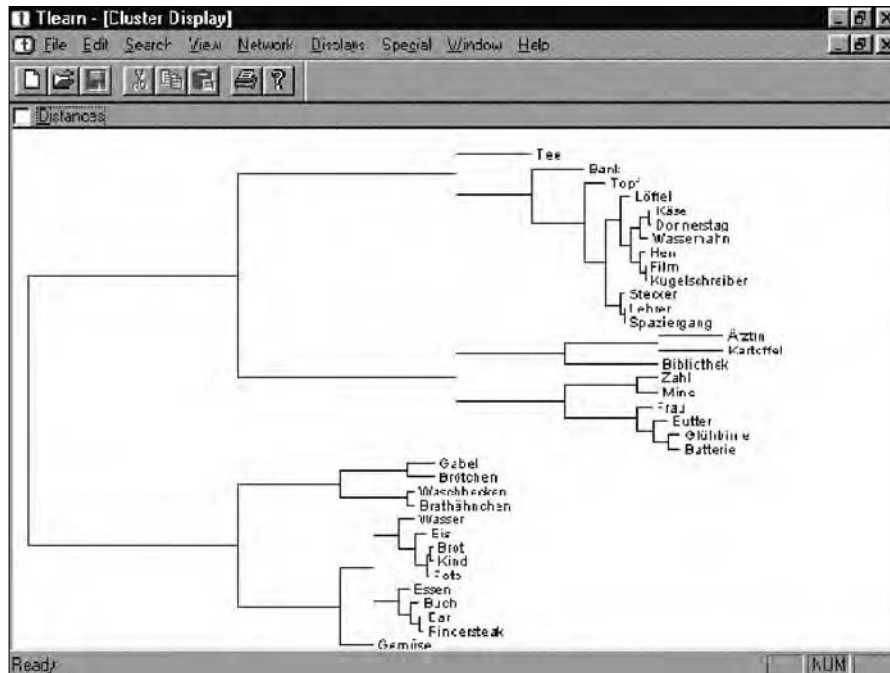


Figure 17. Cluster display

network performance and study data, the simulation has clearly demonstrated that the formal structure of German nouns offers a considerable amount of information on their gender class, and that these formal regularities can well be exploited by the learner solely on the basis of distributed co-occurrence patterns.

## 6. Conclusion

The analysis of the described studies on gender assignment of L2 learners of German shows that even in the early stages of instructed second language acquisition, learners are capable of extracting some of the semantic and formal gender indicators from the input and are able to apply these gender cues when categorizing previously unknown items. The most important factor for the establishment of a learner-internal system of gender assignment patterns obviously is the frequency of nouns conveying a specific gender cue in the learner's input – irrespective of the existence of 'exceptional' assignments – rather than the reliability of this gender indicator.



This finding casts at least some doubt on the widespread practice in language classrooms to primarily teach the ‘waterproof’ gender regularities (such as *-lein* and *-chen* neu, or *-ung* and *-heit/-keit* fem) as gender rules especially in beginners’ classes, as the input frequencies of these regularities are relatively low<sup>38</sup>. The important influence of input frequencies for the deduction of gender regularities especially in the earlier phases of language learning is in concordance with McDonald’s (1989) cue-category-mapping concept, which claims that input frequency and availability of the individual gender indicators are more important for establishing cue validity in early acquisitional stages, while reliability and conflict validity figure larger in later stages<sup>39</sup>. Emphasizing only the most reliable gender rule in the early stages of German L2 learning might thus work counter-productive with respect to the development of a learner-internal gender assignment system, as it would confound the sequence of acquisitional stages. Furthermore, the cost-gain relation of explicitly teaching these ‘waterproof’ regularities as gender rules in beginners’ classes is quite questionable, since students would have to memorize a number of rules with very few opportunities to actually apply them.<sup>40</sup> Instead, it seems much more in line with the findings described above to rather alert students for the general correlations between formal word properties and gender class, even though most of these regularities do have a number of exceptions.

A comparison of type and token frequencies of nouns in the learners’ input and the results of Study 1 furthermore reveals a close correspondence between type frequencies for nouns with a certain gender regularity and accuracy rates for the respective assignments on the one hand, and a correlation of high token frequencies with shorter latencies on the other. This finding corresponds to Bybee’s (1995) Network Model which states that type frequencies are crucial for establishing morphological schemes because they strengthen the internal representation and thus the productivity of a morphological pattern. This result is insofar interesting for the GFL classroom, as it suggests that adequate input enhancement and grouping of gender cue related nouns should assist students to build up gender assignment patterns from early stages of language learning on.

A connectionist simulation of the study data finally offers an explanation of how students scan the formal properties of nouns in the input and consequently develop a learner-internal system of formal gender assignment cues. In their capacity of solving pattern recognition tasks on the basis of incremental algorithmic operations, connectionist networks thus illustrate how learners can build internal representations of the target language structure solely by extracting co-occurring patterns from the input.

**Appendix A**

Type and token frequencies for nouns conveying the various formal gender indicators (students' input and items of Study 1)

Formal gender regularity	Input		Study 1	
	<i>Types</i>	<i>Tokens</i>	<i>Types</i>	<i>Tokens</i>
-a	6	30	1	1
-at	1	1	1	1
-chen	4	8	3	3
-e	85	266	9	9
-el	6	6	1	1
-en	10	37	2	2
-er	40	135	8	8
Ge-	8	36	2	2
-ik	6	13	1	1
-in	14	17	2	2
-keit	1	1	1	1
-ment	1	2	1	1
-o	7	31	4	4
-tion	3	14	1	1
-um	1	1	2	2
-ung	18	55	2	2
-us	1	2	2	2
Ex-e	13	63	2	2
Ex-el	8	17	3	3
Ex-en	11	19	6	6
Ex-er	28	113	2	2

**Appendix B**

Type and token frequencies for nouns conveying the various types of gender regularities (students' input and items of Study 1)

Type of gender regularity	Input		Study 1	
	<i>Types</i>	<i>Tokens</i>	<i>Types</i>	<i>Tokens</i>
Seman. reg.	82	287	8	8
Formal reg.	209	640	43	43
Monosyll.	53	192	10	10
Nomin. verbs	6	11	2	2
Analogies	106	217	8	8
Excep. formal	67	240	13	13
Excep. monosyll.	50	361	16	16

## Appendix C

Comparison of gender assignment accuracy in Study 1 and Network performance

Test item	Gender class	Accuracy of assignment in Study 1	Activation of network output units		
			<i>masc</i>	<i>neu</i>	<i>fem</i>
Film	masc	30.8%	0.984	0.065	0.004
Herr	masc	63.7%	0.996	0.076	0
Tee	masc	25.8%	0.602	0.009	0.019
Topf	masc	54.9%	0.887	0.092	0.025
Bank	fem	26.4%	0.815	0.261	0.005
Frau	fem	71.4%	0.004	0.068	0.964
Bar	fem	25.3%	0.14	0.977	0
Zahl	fem	38.5%	0.058	0.096	0.57
Brot	neu	40.7%	0.016	0.988	0.003
Kind	neu	49.5%	0.011	0.979	0.014
Eis	neu	59.3%	0.06	0.982	0
Buch	neu	40.7%	0.152	0.94	0.002
Käse	masc	30.8%	0.936	0.016	0.01
Stecker	masc	71.7%	0.983	0.124	0.001
Löffel	masc	42.9%	0.95	0.049	0.035
Lehrer	masc	60.4%	0.973	0.108	0.002
Butter	fem	46.2%	0.08	0.076	0.951
Mine	fem	70.3%	0.022	0.067	0.669
Gabel	fem	39.6%	0.789	0.898	0
Ärztin	fem	50.5%	0.485	0.003	0.737
Wasser	neu	36.3%	0.055	0.915	0.016
Foto	neu	76.9%	0.003	0.996	0.003
Essen	neu	35.2%	0.191	0.966	0.003
Brötchen	neu	30.8%	0.661	0.932	0
Donnerstag	masc	52.7%	0.976	0.024	0.01
Wasserhahn	masc	41.8%	1	0	0
Kugelschreiber	masc	69.2%	0.985	0.055	0.001
Spaziergang	masc	56.0%	0.974	0.106	0.003
Kartoffel	fem	27.5%	0.465	0.13	0.454
Glühbirne	fem	67.0%	0.101	0	0.958
Bibliothek	fem	29.7%	0.285	0.314	0.27
Batterie	fem	68.1%	0.039	0.008	0.96
Rindersteak	neu	34.1%	0.119	0.977	0.005
Waschbecken	neu	37.4%	0.425	0.956	0.001
Brathähnchen	neu	26.4%	0.391	0.98	0
Gemüse	neu	22.0%	0.025	0.932	0.149
Brief	masc	44.0%	0.984	0.22	0
Kopf	masc	59.3%	0.847	0.069	0.012
Tür	fem	36.3%	0.353	0.253	0.016

Hand	fem	37.4%	0.592	0.788	0
Spiel	neu	22.0%	0.61	0.532	0.022
Bein	neu	48.4%	0.599	0.28	0.018
Vater	masc	57.1%	0.634	0.611	0.002
Schnupfen	masc	42.9%	0.577	0.003	0.853
Freundin	fem	57.1%	0.961	0.015	0.11
Grippe	fem	64.8%	0.006	0.085	0.882
Büro	neu	41.8%	0.43	0.962	0
Gesicht	neu	34.1%	0.498	0.047	0.098
Supermarkt	masc	46.2%	0.702	0.099	0.013
Kindergarten	masc	44.0%	0.933	0.041	0.002
Besserung	fem	25.3%	0.449	0.024	0.291
Haltestelle	fem	56.0%	0.249	0.008	0.693
Skifahren	neu	23.1%	0.902	0.066	0.006
Medikament	neu	29.7%	0.322	0.957	0.001
Weg	masc	34.1%	0.793	0.837	0
Turm	masc	45.1%	0.169	0.524	0.075
Kuh	fem	29.7%	0.812	0.429	0.001
Wahl	fem	16.5%	0.285	0.126	0.029
Rad	neu	34.1%	0	0.841	0.766
Huhn	neu	29.7%	0.949	0	0.02
Bürger	masc	69.2%	0.933	0.061	0.007
Kaktus	masc	36.6%	0.998	0.014	0
Fähre	fem	60.4%	0.059	0.004	0.985
Technik	fem	20.9%	0.99	0.032	0.003
Zentrum	neu	34.1%	0.957	0.027	0.037
Motto	neu	50.5%	0.312	0.516	0.014
Campingurlaub	masc	50.5%	0.984	0.144	0.001
Reiseführer	masc	65.9%	0.605	0.502	0.006
Information	fem	24.2%	0.107	0.509	0.133
Langeweile	fem	65.9%	0.67	0.002	0.356
Lebensmittel	neu	24.2%	0.522	0.395	0.096
Jubiläum	neu	28.6%	0.998	0.022	0
Sosch	masc	28.6%	0.992	0.324	0
Krap	masc	42.9%	0.001	0	0.999
Pila	fem	33.0%	0.005	0.945	0.212
Wote	fem	59.3%	0.007	0.091	0.895
Dilo	neu	48.4%	0.592	0.653	0
Lett	neu	42.9%	0.084	0.989	0.001
Nänter	masc	58.2%	0.893	0.047	0.063
Klapptus	masc	31.9%	0.955	0.083	0.006
Reike	fem	67.0%	0.412	0.026	0.431
Griehlung	fem	24.2%	0.411	0.002	0.789
Mosat	neu	36.3%	0.655	0.037	0.2
Pläckchen	neu	28.6%	0.162	0.992	0.001
Bitzefisch	masc	44.0%	0.922	0.134	0.006
Danborenker	masc	75.8%	0.963	0.363	0.001
Faltonbar	fem	25.3%	0.902	0.057	0.038
Näseperkeit	fem	46.2%	0.169	0.009	0.625
Tinkerbrot	neu	36.3%	0.83	0.613	0.002
Schenpalarum	neu	20.9%	0.926	0.001	0.074

## Notes

1. Excerpt from an interview with an eight-year-old Russian speaking learner of German as a second language after eleven months of living in Germany, from: Wegener (1995: 6–7).
2. For a detailed overview see the two special issues “Processing of Grammatical Gender I +II” of the *Journal of Psycholinguistic Research* 28 (5)(6) 1999.
3. All words were embedded in the sentence: “*J’ai vu un(e)* [resp. *de* in the neutral condition] *joli(e)* \_\_\_\_\_ *samedi*.” [I saw a / some nice \_\_\_\_\_ last Saturday.]
4. Aufderstraße, Hartmut, Heiko Bock, Mechthild Gerdes, Jutta Müller and Helmut Müller (1992). *Themen Neu Kursbuch 1*. Ismaning: Hueber
5. Students were, however, made aware of the importance of the gender class of each noun for the production of correct noun phrases. Nevertheless, they were not deliberately instructed to memorize noun gender, and gender knowledge was not tested in form of gender assignment to individual nouns in control tests.
6. I am grateful to Katsuo Tamaoka for his most valuable help with the preparation and carrying-out of the study.
7. For a more detailed discussion of this aspect and a comparison of overall German achievement and gender knowledge see Menzel (2004: 191–205).
8. Since the items were presented in written form, morphological and phonological regularities will not be differentiated. Furthermore, root endings, suffixes and pseudo suffixes are not distinguished, as it is rather improbable that students are aware of this distinction let alone employ it for gender assignment.
9. This gender regularity is based upon the tendency for monosyllables to take masculine gender (64% of all monosyllabic nouns listed in Duden <sup>14</sup>1954 are classified as masculines). The various formal gender regularities applying to monosyllables (Köpcke 1982) have not been taken into consideration here.
10. This type of gender regularities includes compounds following the last member rule as well as rhyme analogies between nonce words and real nouns known to the subjects.
11. All items in this group are nouns known to all subjects.
12. See Menzel (2004: chap. 3.2.2) for a detailed description of these results.
13. For a complete list of all results see Menzel (2004: 129).
14. About 90% of the 15.000 nouns ending in *-e* take feminine gender (Mills 1986: 33). See appendix A for an overview of type and token frequencies for nouns with the various formal gender indicators (students’ input and items of Study 1).
15. None of the participants had any contact with the German language outside of the classroom setting, and all classes used the same textbook.
16. It should be noted, however, that although this way of calculating the subjects’ input certainly offers a more accurate approximation of the students’ real GFL learning history as e.g. an estimation based upon general word frequency lists

would, these *input* figures nevertheless should not be confounded with actual *intake*. Repetition of certain textbook passages in class, as well as the intensity of home preparation and students' attention vary from subject to subject and throughout the school year.

17. I am grateful to Ursula Menzel for her assistance with this time-consuming task.
18. Note that the gender assignment study included 30 items from each gender class.
19. Latencies are calculated from both correct and incorrect answers. Correct assignments (2167 ms) yielded significantly shorter latencies than incorrect answers (2349 ms):  $t(90) = -6.8$ ;  $p < 0.000$ . Overall correlation between accuracy and shorter latencies is extremely high ( $r = 0.971$ ;  $n = 91$ ;  $p < 0.000$ ), i.e., there is very little, if any, risk of speed-accuracy trade-off.
20. See appendix B for an overview of type and token frequencies for nouns with the corresponding types of gender regularities (students' input and items of Study 1).
21. For the sake of clearness, these results are presented in two graphs. See appendix A for an overview of type and token frequencies for nouns with the depicted formal gender indicators.
22. See Menzel (2004: chap. 3.5.2) for a discussion of these deviant results.
23. Due to space limitations, not all results of this study can be presented here. For a detailed overview see Menzel (2004: 187–190).
24. The latency for *-en* deviates significantly from the average reaction time for all subjects over all items ( $t(90) = -5.36$ ;  $p < 0.000$ ), whereas the result for *Ex-en* just misses the significance level ( $t(90) = -1.84$ ;  $p = 0.068$ ).
25. For a detailed comparison between model performance and the empirical data see MacWhinney et al. 1989: 268–269).
26. For a detailed discussion of this problem see McLeod et al. 1998; Rolls and Treves 1998; Plunkett and Elman 1997.
27. The mathematical basis for this operation is the generalized Delta-Rule:  $\Delta w = \alpha(\lambda - w)$ .
28. See McLeod et al. (1998) and Rolls and Treves (1998) for an introductory overview of various modelling approaches.
29. I would like to thank Kim Plunkett and Michael Thomas for their valuable advice in model construction.
30. See chap. 4.2 for an account of the estimation of the students' nominal input.
31. Alternatives to solving the problem of different word lengths in input coding could have been to choose items with a corresponding number of letters only, or to confine coding to the first and the last syllable of each item (as done by MacWhinney et al. 1989, in their simulation).
32. There are other types of connectionist networks that are able to self-determine the optimal number of hidden units by starting out with a relatively small number of hidden units and recruiting more in case no satisfying solution has been found.
33. For a detailed description of the mathematical operations employed in calculating the RMS error rate see Plunkett and Elman (1997: 47).

34. After ten runs through the training set the RMS error amounted to 0.25 and after 25 runs to 0.1.
35. Note that although the RMS error was reduced significantly after five additional runs through the training set, the network then classified the testitem *Gemüse* [vegetables] as feminine, while the general output activation for the item *Tee* [tea] was considerably weakened.
36. These options will be introduced only briefly here, for a completer account see e.g. Plunkett and Elman (1997).
37. For the sake of clearness, Figure 17 only includes 36 of the 90 test items.
38. Wegera (1997: 65) e.g. lists *Ausnahmslosigkeit* (exceptionlessness) as the most crucial criterion for gender rules in his didactical reduction: “Als Lernregeln sollten nur Regularitäten herangezogen werden, die möglichst ausnahmslos gelten und nicht nur Tendenzen darstellen. [...] Bereits wenige Ausnahmen zu einer Regel schmälern ihren Nutzen als Lernregel erheblich.” [Only regularities with possibly no exceptions should be drawn upon as learning rules, but not those that merely depict tendencies. [...] Only a small number of exceptions already considerably reduces the use of [gender] learning rules.]
39. This sequence in the acquisition of cue properties can also be found in connectionist networks with longer training experience. McDonald (1989: 392–396) analyzed Taraban’s et al. (1989) connectionist simulations of the acquisitional process of the German definite article declension paradigm with respect to the development of cue properties and discovered a very similar pattern: “In the early stages of learning, overall validity is the best predictor of connection strengths. Later, cue reliability is briefly the best predictor, but in the last stages of learning, conflict and cooperation relations [...] are the best predictors of network weights.” (McDonald 1989: 394)
40. Compare the list of type and token frequencies for the respective nouns in appendix A.

## References

- Altmann, G. and V. Raettig  
1973     Genus und Wortauslaut im Deutschen. In *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 2: 297–303.
- Andersen, Roger W.  
1984     What’s gender good for, anyway?. In *Second Languages: A Cross-Linguistic Perspective*, Roger W. Andersen (ed.), 77–99. Rowley: Newbury House.
- Aufderstraße, Hartmut, Heiko Bock, Mechthild Gerdes, Jutta Müller and Helmut Müller  
1992     *Themen Neu Kursbuch 1*. Ismaning: Hueber.

- Bates, Elizabeth and Brian MacWhinney  
 1989 Functionalism and the Competition Model. In *The Crosslinguistic Study of Sentence Processing*, Brian MacWhinney, and Elizabeth Bates (eds.), 3–73. Cambridge: Cambridge University Press.
- Bates, Elizabeth and Brian MacWhinney  
 1987 Competition, Variation, and Language Learning. In *Mechanisms of Language Acquisition*, Brian MacWhinney (ed.), 157–193. Hillsdale: Erlbaum.
- Bybee, Joan  
 1988 Morphology as lexical organization. In *Theoretical Morphology*, Michael Hammond and Michael Noonan (eds.), 119–141. San Diego.  
 1995 Regular morphology and the lexicon. In *Language and Cognitive Processes* 10: 425–455.
- Corbett, Greville  
 1991 *Gender*. Cambridge: Cambridge University Press.
- Duden-Rechtschreibung  
 1954 Mannheim: Dudenverlag.
- Eisenberg, Peter  
 1994 *Grundriß der deutschen Grammatik*. Stuttgart: Metzler.
- Feldman, Jerome A.  
 1981 A connectionist model of visual memory. In *Parallel Models of Associative Memory*, Geoffrey A. Hinton, and John R. Anderson (eds.), 49–81. Hillsdale: Erlbaum.
- Friederici, Angela D. and Thomas Jacobsen  
 1999 Processing grammatical gender during language comprehension. In *Journal of Psycholinguistic Research* 28 (5): 467–484.
- Gregor, Bernd  
 1983 *Genuszuordnung*. (Linguistische Arbeiten 129). Tübingen: Niemeyer.
- Grosjean, François, Jean-Yves Dommergues, Etienne Cornu, Delphine Guillemon and Carole Besson  
 1994 The gender-marking effect in spoken word recognition. In *Perception and Psychophysics* 56: 590–598.
- Hagoort, Peter and Colin M. Brown  
 1999 Gender electrified: ERP evidence on the syntactic nature of gender processing. In *Journal of Psycholinguistic Research* 28(6): 715–728.
- Hoberg, Ursula  
 1999 Das Genus des Nomens. In *Grammatik des Deutschen im europäischen Vergleich*. Unpublished manuscript.
- Jacobsen, Thomas  
 1999 Effects of grammatical gender on picture and word naming: Evidence from German. In *Journal of Psycholinguistic Research* 28(5): 499–514.
- Karmiloff-Smith, Annette  
 1979 *A Functional Approach to Child Language*. Cambridge: Cambridge University Press.



- Köpcke, Klaus-Michael  
1982 *Untersuchungen zum Genusssystem der deutschen Gegenwartssprache*. Tübingen: Niemeyer.
- Köpcke, Klaus-Michael and David Zubin  
1983 Die kognitive Organisation der Genuszuweisung zu den einsilbigen Nomen der deutschen Gegenwartssprache. In *Zeitschrift für germanistische Linguistik* 11: 166–182.  
1984 Sechs Prinzipien für die Genuszuweisung im Deutschen: Ein Beitrag zur natürlichen Klassifikation. In *Linguistische Berichte* 93: 26–50.
- Lang, Adrienne  
1976 The semantic base of gender in German. In *Lingua* 40: 55–68.
- MacWhinney, Brian  
1989 Competition and connectionism: In *The Crosslinguistic Study of Sentence Processing*, Brian MacWhinney and Elizabeth Bates (eds.), 422–457. Cambridge: Cambridge University Press.  
1978 *The Acquisition of Morphophonology*. (Monographs of the Society for Research in Child Development 43, 1/2. Chicago: University of Chicago Press.  
1987 The Competition Model. In *Mechanisms of Language Acquisition*, Brian Mac Whinney (ed.), 249–308. Hillsdale: Erlbaum.
- Maratsos, Michael and Anne Chalkley  
1980 The internal language of children's syntax: The ontogenesis and representation of syntactic categories. In *Children's Language*, Vol II, Keith E. Nelson (ed.), 127–214. New York: Gardner.
- McDonald, Janet  
1989 The acquisition of cue-category mappings. In *The Crosslinguistic Study of Sentence Processing*, Brian MacWhinney and Elizabeth Bates (eds.), 375–396. Cambridge: Cambridge University Press.
- McLeod, Peter, Kim Plunkett and Edmund T. Rolls  
1998 *Introduction to Connectionist Modelling of Cognitive Processes*. Oxford: Oxford University Press.
- Meinert, Roland  
1989 *Die deutsche Deklination und ihre didaktischen Probleme*. München: iudicium.
- Menzel, Barbara  
2004 Genuszuweisung im DaF-Erwerb: Psycholinguistische Prozesse und didaktische Implikationen. Berlin: Weisensee.
- Mills, Anne E.  
1985 The acquisition of German. In Dan I. Slobin (ed.), *The Crosslinguistic Study of Language Acquisition, Volume 1: The Data*, 141–254. Hillsdale: Erlbaum.  
1986 *The Acquisition of Gender*. Berlin Springer.
- Minsky, Marvin L. and Seymour Papert  
1969 *Perceptrons*. Cambridge: Cambridge University Press.

- Pfaff, Carol W.  
1992 The issue of grammaticalization in early German Second Language. In *Studies in Second Language Acquisition* 14: 273–296.
- Plunkett, Kim and Jeffrey L. Elman  
1997 *Exercises in Rethinking Innateness*. Cambridge: MIT Press.
- Rolls, Edmund T. and Alessandro Treves  
1998 *Neural Networks and Brain Function*. Oxford: Oxford University Press.
- Rumelhart, David E. and James D. McClelland (eds.)  
1986 *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Cambridge.
- Schmidt, Rainer  
1986 Was weiß der Artikel vom Hauptwort?. In *Zeitschrift für experimentelle und angewandte Psychologie* XXXIII (1): 150–163.
- Schriefers, Herbert  
1993 Syntactic processes in the production of noun phrases. In *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19 (4): 841–850.
- Sokolik, M.E. and Michael E. Smith  
1992 Assignment of gender to French nouns in primary and secondary language: A connectionist model. In *Second Language Research* 8 (1): 39–58.
- Strong, Mark  
1976 German noun gender: A matter of memorization? In *Die Unterrichtspraxis* 9: 74–77.
- Taraban, Roman and Vera Kempe  
1999 Gender processing in native and nonnative Russian speakers. In *Applied Psycholinguistics* 20: 119–148.
- Taraban, Roman, Janet L. McDonald and Brian MacWhinney  
1989 Category learning in a connectionist model: Learning to decline the German definite article. In *Linguistic Categorization*, Roberta Corrigan, Fred Eckmann, and Michael Noonan (eds.), 164–193. Amsterdam: Benjamins.
- van Berkum, Jos  
1996 *The psycholinguistics of grammatical gender*. Nijmegen: Nijmegen University Press.
- Wegener, Heide  
1995 Das Genus im DaZ-Erwerb. Beobachtungen an Kindern aus Polen, Rußland und der Türkei. In *Fremde Sprache Deutsch*, Brigitte Handwerker (ed.), 1–24. Tübingen: Narr.  
2000 German gender in children's second language acquisition. In *Gender in Grammar and Cognition*, Barbara Unterbeck, and Matti Rissanen (eds.), 511–544. Berlin: Mouton de Gruyter.

Wegera, Klaus-Peter

1997 *Das Genus*. München: iudicium.

Werner, Otmar

1975 Zum Genus im Deutschen. In *Deutsche Sprache* 3: 35–58.



## **Language analytic ability and oral production in a second language: Is there a connection?**

*Leila Ranta*

A relatively large body of research has established that learners who have greater levels of 'language analytic ability' achieve higher scores on measures of L2 learning (Carroll 1962; Ranta 2002; Skehan 1986). A few studies have also found a relationship between analytic ability and global ratings of performance on oral tasks (Ehrman & Oxford 1995; Horwitz 1987; Harley & Hart 2002). However, from a cognitive processing perspective, it is predicted that analytic ability will be associated with grammatical development or accuracy rather than with fluency in speaking (Skehan 2002). The study reported here is a test of this prediction. Oral production data, which had been collected in a previous study (Lightbown & Spada 1997), were analysed with respect to both grammatical development and to fluency. Learners were classified as belonging to stages of interlanguage development for question forms and for possessive determiners. Fluency was assessed using six measures including speech rate, pausing, and self-repair variables. The learners were francophone children studying in an intensive ESL program in Quebec, Canada. The ESL program promoted oral interpersonal communication skills rather than formal accuracy or academic proficiency in English. Two groups of learners (n=15 each) were formed based on their performance on an L1 and a L2 metalinguistic task (Ranta 1998, 2002b). The analytic learners were found to be significantly more advanced than the less analytic learners with respect to the production of possessive determiners; the difference between the two groups for the question form analysis was not significant, although the trend was in the same direction. In contrast, no significant differences were found between the groups with respect to five of the six fluency measures. These results suggest that the fluency-oriented nature of L2 instruction in this program led to relatively uniform levels of fluency among learners who differed in analytic ability but that the lack of form-focussed instruction may have disadvantaged less analytic learners in terms of grammatical development.

This paper reports on the findings from a study of francophone learners in an intensive English as a second language (ESL) program in Quebec, Canada. Although a fairly narrow topic is addressed, it is embedded in a larger program of research conducted in this unique instructional context under the leadership of Patsy Lightbown and Nina Spada and in collaboration with other colleagues and graduate students (see overview in Lightbown & Spada 1994). Intensive ESL classrooms offer the opportunity to observe learners make great gains in second language (L2) learning over a short period of time. In addition, they are instructional environments where a strong version of communicative language teaching has been successfully implemented. The research program in intensive ESL began with a descriptive study involving classroom observation and language testing in many different classrooms; this provided a rich database to explore the relationship between characteristics of teaching behaviour and L2 learning outcomes (Lightbown & Spada 1990; Spada 1990). Lightbown and Spada concluded from this phase of the research that “accuracy, fluency and overall communicative skills are probably best developed through instruction that is primarily meaning-based but in which guidance is provided through timely form-focus activities and correction in context” (1990: 433). This became the hypothesis that motivated a number of teaching experiments where learners were exposed to form-focussed instructional activities relating to a particular target structure of the L2 during their otherwise communicative “bath” (e.g., L. White, Spada, Lightbown, & Ranta 1991; J. White 1996, 1998; Spada, Lightbown, & J. White, this volume; J. White & Ranta 2002).

My involvement in data collection during many of these studies sparked an interest in the differences among individual learners, in particular, the role of learners’ analytic abilities in the L2 acquisition process (Ranta 1998). As I searched the literature, I found that although there were strong opinions that analytic ability is not relevant to L2 learning in a communicative setting (e.g., Cook 1996), very little research had been conducted in classrooms that could unambiguously be characterized as communicative. I therefore undertook a study of the relationship between learners’ analytic ability and their performance on a variety of L2 proficiency measures (Ranta 1998, 2002b). The results from a cluster analysis revealed that the strongest group of learners demonstrated the ability to handle a variety of analytic and decontextualized tasks despite the fact that the instructional content did not promote the development of such skills. These learners also scored highest on an L1 metalinguistic task. In this paper, I will further explore the impact of learners’ analytic ability on L2 acquisition by examining oral production data from these same learners.

### **Language analytic ability**

What makes some individuals better language learners than others? One answer to this question is that success/failure is due to differences in the learners' foreign language aptitude, a construct which consists of three component abilities – auditory ability, language analytic ability, and memory abilities (Skehan 1989, 1998, 2002). With respect to analytic ability, the model of foreign language aptitude proposed by Carroll (1962, 1981) distinguished between grammatical sensitivity defined as the ability “to recognize the grammatical functions of words in sentences” and inductive language learning ability which referred to “the ability to infer or induce the rules governing a set of language materials, given samples of language materials that permit such inferences” (Carroll 1981: 105). Skehan (1989) later collapsed the two components into one called “language analytic ability” and defined it as “the capacity to infer rules of language and make linguistic generalizations or extrapolations” (Skehan 1998: 207). In a recent paper, Skehan (2002) has found it useful once again to separate the two components in order to address different aspects of the acquisition process (see below). Chapelle and Green (1992) also argue for two types of analytic ability so as to distinguish between “fluid”, that is innate abilities, and “crystallized” or learned abilities. In contrast to this view, I would like to stress the underlying similarity between fluid and crystallized abilities relating to language analysis.

Language analytic ability as defined above covers both implicit processes found in the child's acquisition of syntactic structures (e.g., Wong Fillmore 1979) and the explicit processes that are reflected in the kinds of questions that L2 teachers of adults routinely face. Some scholars distinguish between the type of linguistic analysis and hypothesis-testing of the child from similar activities performed by the professional linguist, but there is strong evidence that implicit analysis and metalinguistic ability are related to each other. The follow-up study by Skehan (1989, 1990) of some of the children who had participated in the Bristol L1 development study (Wells 1985) revealed moderate correlations between measures of L1 development and scores on the Language Analysis subtest of the Pimsleur Language Aptitude Battery (PLAB: Pimsleur 1966). As I have argued elsewhere (Ranta 1998, 2002b), the concept of language analytic ability overlaps with metalinguistic ability since aptitude is always measured with some kind of metalinguistic task. At the heart of both concepts is the ability of an individual to focus on the structural properties of linguistic utterances rather than on their meaning<sup>1</sup>.

Analytic ability consistently emerged as a predictor in studies of successful language learning that took place in traditional or audiolingual classrooms (e.g., Carroll 1962, 1981; Gardner & Lambert 1972). Despite its strong track record, the analytic component of aptitude lost prominence once language teaching methodologies changed from emphasizing the formal aspects of the target language to emphasizing how to communicate in it. This led many educators and scholars to assume that learners' analytic abilities were no longer relevant to L2 learning. After all, if explicit grammatical explanation and exercises are not a focus of teaching and testing, why should the ability to perform metalinguistic analysis be important? Explicit statements concerning the limited relevance of foreign language aptitude, including analytic ability, can be found in the writings of SLA scholars such as Krashen (1981) and Cook (1996). The anti-aptitude view is also implicit in publications aimed at pre-service teachers where aptitude is not mentioned at all (e.g., Hadley 2001; Shrum & Glisan 2000), or is relegated to peripheral status as one of many learning styles/strategies (Scarcella & Oxford 1992). The research evidence, however, suggests that aptitude, including language analytic ability, does have an impact on L2 learning outcomes even in communicative teaching contexts (see review in Sawyer & Ranta 2001). Furthermore, recent developments in SLA theory are completely consistent with the role for analytic ability in the acquisition process in naturalistic settings. In the following section, we will consider the framework provided by Skehan (2002).

### **Information-processing models and aptitude**

Although a number of information-processing approaches can be found in the SLA literature (e.g., Bialystok 1994; DeKeyser 1998; McLaughlin & Heredia 1996), the processing model outlined by Skehan (2002) offers some clear advantages. For one thing, the model accounts for learner knowledge that is derived from implicit input-processing as well as from explicit metalinguistic instruction. It also includes a description of a series of output-processing stages that allows a smoother integration of fluency development into the acquisition process. Most importantly for the present discussion, Skehan proposes how different components of aptitude (including those measured by aptitude tests and ones that are not) are related to specific stages in the acquisition process, thus offering specific hypotheses for future research. Skehan's model, which is outlined in Table 1, describes the stages of L2 learning as follows: noticing (Stage 1), patterning (Stages 2–5),



controlling (Stages 6–8), and lexicalising (Stage 9). As can be seen from the table, the language analytic components (i.e., grammatical sensitivity and inductive language learning) are posited to be relevant to the patterning stages in which input that has been noticed is processed for meaning and analyzed structurally, leading to generalizations, hypothesis-testing, and restructuring. In essence, language analytic ability is responsible for driving grammar development forward. From this theoretical framework, one would clearly expect learners with strengths in language analytic ability to acquire L2 grammatical structures more rapidly, and perhaps more extensively than those with a different aptitude profile. In contrast, one would not necessarily expect analytic learners to have strengths in fluency since this calls upon other components of aptitude.

*Table 1.* Model of SLA processing and L2 aptitude adapted from Skehan 2002: 88–90

<b>SLA Processing Stage</b>	<b>Nature of Stage</b>	<b>Potential Aptitude Components</b>
1. noticing	learner directs attention to some aspects of the language system, or is led to direct attention in this way	auditory segmentation attention management working memory phonemic coding
2. pattern identification	the learner makes a hypothesis or generalization based on a perceived pattern or regularity	fast analysis/working memory grammatical sensitivity
3. extending	the learner extends the domain of the hypothesis	inductive language learning ability
4. complexifying	the learner apprehends the limitations of the identified pattern, and restructures it, as new aspects of the target language are noticed	grammatical sensitivity inductive language learning ability
5. integrating	the learner integrates the new form into a larger structure	restructuring capacity
6. becoming accurate	learner becomes able to use the new form without making errors, although this use may be slow and effortful	automatisation proceduralisation
7. creating a repertoire	the new form can be accessed at appropriate places	retrieval processes
8. achieving fluency	the new form now used with reasonable speed and accuracy	automatising, proceduralisation
9. lexicalizing	learner produces the new form as a lexicalized element in addition to rule-based version	memory chunking retrieval processes

**Language analytic ability in communicative classrooms**

Given the above conception of the acquisition process and of language analytic ability as a driver of interlanguage development in naturalistic SLA, we would expect that language analytic ability will advantage learners in oral production when grammatical development is the focus but not when fluency is being evaluated. Unfortunately, empirical evidence bearing on this issue is very limited.<sup>2</sup> Two studies carried out in French immersion classes in Canada included aptitude tests and oral production measures as part of a larger assessment battery (Tucker et al., 1976; Harley & Hart 1997). Tucker et al. had native speakers listen to audio-recordings of immersion students and rate them on five separate scales (comprehensibility, pronunciation, vocabulary, grammar, amount of communication). Harley and Hart (1997) used a sentence repetition task and an oral cartoon description task which was analyzed in terms of the number of lexical verbs used. In neither study did they find a significant relationship between language analytic ability and the oral production measures used. A related study by Harley and Hart (2002) investigated the impact of language analytic ability among Canadian high school students who were experiencing their first intensive exposure to naturalistic French. The sentence repetition task was the only oral measure administered. They found a significant correlation between scores on the Language Analysis subtest of the PLAB and learners' ability to repeat exactly what they heard. While suggestive of a relationship between language analysis and oral production, it is difficult to know just how performance on this task relates to accuracy and fluency in spontaneous oral production.

Two studies that found significant correlations (around  $r = .4$ ) between grammatical sensitivity and measures of "free" oral production are Horwitz (1987) and Ehrman and Oxford (1995). In both studies analytic ability was measured using Words in Sentences. Horwitz rated learners' oral productions on three tasks using a five-point scale; the criteria for communicative effectiveness included amount of communication, fluency, ability to incorporate interviewer's cues, complexity, and length of structures used. Ehrman and Oxford used the ACTFL (American Council for the Teaching of Foreign Languages) rating scales for speaking which define accuracy as "the acceptability, quality and precision of the message conveyed" (ACTFL testers' handbook by Buck, Byrnes, & Thompson 1989 cited in Hadley 1993: 17). According to the handbook, the criteria to be considered when assessing accuracy are fluency, pronunciation, grammar, vocabulary, pragmatic competence, and sociolinguistic competence. Thus, it appears that

Horwitz and Ehrman and Oxford's studies provide evidence that language analytic ability is related to assessments of oral production that reflect overall proficiency, but because the holistic rating scales conflated grammatical accuracy and fluency (along with other criteria), we do not know how analytic ability influences fluency separately from accuracy.

### **The development of fluency**

In the language teaching literature, the term "fluency" is used in two distinct ways: in some contexts, fluency refers to overall oral proficiency in a L2, in others, it describes speech that is rapid and smooth (Koponen & Riggenbach 2000; Lennon 2000). Researchers interested in fluency as smoothness (referred to by Lennon 2000: 25 as "lower-order fluency") typically assess fluency through temporal variables such as speech rate and frequency of pausing, and dysfluency markers such as hesitations, repetitions, and self-repairs (i.e., self-corrections) (Lennon 1990; Raupach 1980; Riggenbach 1991; see review in Koponen & Riggenbach 2000). The controlling and lexicalising stages (7–9) in Skehan's model are concerned with the development of lower-level fluency.

Skehan (2002) proposes that individual differences in the memory component of aptitude, rather than analytic ability, will lead to differences in fluency among learners. The case study of CJ illustrates his claim (Obler 1989). CJ was reported to be able to learn a number of languages rapidly; his scores on Words in Sentences were merely average but his memory for verbal material was phenomenal. In earlier work, Skehan (1989, 1998) proposed that analysis-oriented learners prioritize accuracy whereas memory-oriented learners prioritize fluency. Some support for this comes from the study by Kormos (1999) who examined the relationship between self-reported concern for accuracy versus fluency among Hungarian university-level learners of English. Using performance on a proficiency measure as a covariate in her analysis, she found that there was a relationship between "speaking style" and speech rate such that the accuracy-oriented learners spoke more slowly than the fluency-oriented learners. The two groups did not, however, differ in terms of overall frequency of self-repairs.<sup>3</sup>

Information-processing models taking a universalist rather than individual differences approach (Skehan 1991) view fluency development as resulting from extensive practice which has allowed the underlying processes involved in speech production to become automatized (Schmidt 1992). Some of the characteristics of an automatized process is that it is

rapid, does not require effort or attention, and is difficult to modify or inhibit. The influential model by Anderson (e.g., 1985) known as Adaptive Control of Thought (ACT)<sup>4</sup> clarifies how new skills are acquired. In the model, practice transforms declarative knowledge, which is knowledge about the world, into procedural knowledge or knowledge how to do things. Initial practice serves to proceduralize declarative knowledge and then further practice leads to gradual automatization (see more detailed descriptions of the model in DeKeyser 1998; Schmidt 1992). Validation of this general model for L2 acquisition comes from DeKeyser (1997) who conducted a lab experiment involving highly controlled practice tasks and an artificial language as the target. Among real-world learners, Towell, Hawkins, and Bazergui (1996) argue that proceduralization of declarative knowledge offers the best explanation for the improvement in fluency they found among 12 British university students of French after a study-abroad experience of six months in France.

The ACT theory and the studies by DeKeyser and Towell et al. do not provide much guidance about the type of practice that leads to fluency. Foster and Skehan in a series of experiments with ESL learners (Foster & Skehan 1996; Skehan & Foster 1999) have explored the effect of task demands on learners' performance. They have shown how characteristics of tasks predispose learners to channel their attention in predictable ways, leading to trade-off effects between fluency, accuracy, and complexity in their production. From a slightly different perspective, Segalowitz (1997, 2000) emphasizes that practice activities must be transfer-appropriate. This means that they involve the activation of "cognitive operations that are likely to be reinstated later when the individual attempts to put the learning into practice" (Segalowitz 2000: 213). With respect to developing oral fluency, learners need to experience practice that is extensive and repetitive in order to build automaticity, but also practice that is genuinely meaningful so as to be transfer-appropriate.

From this brief review of the fluency literature, it is apparent that two competing predictions about the relationship between analytic ability and fluency can be made. If we see individual differences as the source of inter-individual variation, we might expect that the more analytic learners will speak more slowly in order to plan their utterances and will be more prone to self-repair; the non-analytical learners may rely more on their memories and therefore be able to achieve fluent performance more quickly. The alternate prediction is that learning in a second language situation where opportunities to practice are similar for all learners should lead to uniform levels of automaticity.

One difficulty with making predictions on the basis of the existing fluency literature is that coverage has been in most cases limited to learners who were adults and at advanced levels of L2 proficiency. Would the same be necessarily true for younger learners? For example, Kormos (1999) argues on the basis of her research with EFL university students that the frequency of self-repairs does not change with competence. And yet, a developmental pattern in which the frequency of self-repairs increased from kindergarten to grade 2 has been found in English L1 (Evans 1986) and among young L2 learners of Dutch (Verhoeven 1989). Also, we might ask whether speaking style differences found among adults after years of L2 study will be evident in the early stages of L2 learning among younger learners. This study aims to shed some light on the relationship between analytic ability and learners' accuracy and fluency in oral production by addressing the following research questions. Given the communicative learning context of intensive ESL classes:

1. Do learners with greater language analytic ability outperform learners with less analytic ability in terms of interlanguage stage development?
2. Do analytic learners differ from less analytic learners in terms of fluency as measured by commonly used measures of temporal variables and repair phenomena?

## **METHOD**

### **Research context**

This study took place in a French-language school in the province of Quebec, Canada. The participants were grade 6 francophone students who attended an intensive ESL program in the Montreal area (Lightbown & Spada 1994). Unlike ESL instruction in the rest of Canada, French-speaking learners are usually in a context that can be characterized as "English as a foreign language" in that they have relatively little contact with English outside of school, except for English-language media, and are not required to develop academic proficiency in English. This study was carried out in a special school where all students were in grade 6 and all experienced intensive language learning. In this school, five classes began the year with an all-day program in ESL while another five classes studied the grade 6 subject-matter curriculum, delivered entirely in French, the students' L1. At the end of January, in the middle of the school year, the classes switched.

Previous descriptions of teaching in intensive ESL classrooms have emphasized the instructional focus on meaning rather than on form and on student-centered rather than on teacher-centered activities (Lightbown & Spada 1994; Spada 1990). For the present discussion, it is important to consider whether this communicative program provided opportunities for learners to develop their oral fluency in English. Brumfit (1984) has argued that fluency is best developed in classrooms through small-group activities in which learners produce language which they have processed themselves, expressing content that is determined by them, where normal processes of adjustment and negotiation occur, and where the target language is used as a means to a communicative end. On the basis of many different sources of information about intensive ESL classrooms including an in-depth analysis of instructional materials (Weary 1987), formal observations using the COLT observation scheme (Spada & Fröhlich 1995) among a large number of classes (Lightbown & Spada 1990; Spada 1990; Spada & Lightbown 1989), informal observations over a period of 10 years, and responses to a teacher questionnaire (Ranta 1998), it is quite clear that ESL instruction in the intensive classes under study were indeed fluency-oriented.

All of the pedagogical activities had a task-related objective such as to use English to win points, to find something out, plan a skit, solve a problem, or draw a picture. Most of class time was devoted to oral activities such as games, puzzles, surveys, interviews and discussions which were organized around themes relevant to the interests of students such as “all about me, you and my family”, food, and fashion. Oral presentations were a major part of the program, ranging from fairly brief “show and tell” sessions to more elaborate productions like a fashion show where teams worked on their presentations over a period of weeks; these presentations involved choosing a theme, writing descriptions of outfits, selecting background music, decorating the classroom, and performing their show before invited guests.

The inclusion of songs, poems, and tongue twisters offered the only form-oriented practice that regularly occurred throughout the five-month program. Singing was a favourite activity among the students. In the early months, students learned simple songs with accompanying mimes; later, they sang popular songs that offered more complex content and language. For example, I observed a class sing, with great energy and feeling, a Michael Jackson song with the words “And the world will be a better place, for you and for me, and the entire human race”. Singing offered relief from the cognitive demands of speaking in a L2 for the school day but also provided opportunity for enjoyable automatization practice.

**Previous research in the Intensive ESL School**

The oral data discussed in this paper were collected as part of a larger study which focused on describing L2 learning outcomes in the Intensive ESL school (Lightbown & Spada 1997). A number of proficiency measures were administered to the 301 students in the school in September, January and June. These measures included vocabulary recognition and listening comprehension tests, a cloze passage and an error detection task that focused on third person possessive determiners. This task is referred to as the L2 analytic task (see Ranta 1998, 2002b for details).<sup>5</sup> In addition to written tests administered to the class as a whole, a sample of 10 students was randomly selected from each class at the end of their five-month ESL program for participation in two oral production tasks.

My doctoral dissertation (Ranta 1998) focused on the cohort of students who did their ESL program from September to January and their French academic program from February to June (N=150). In addition to the L2 proficiency measures that all students completed, a French error correction task was administered; this is referred to as the L1 analytic task. The results from the written L2 proficiency measures were analyzed using correlations and principal components analysis which revealed a complex relationship among the measures (see details in Ranta 1998, 2002b). The cluster analysis, however, served to clarify the nature of the complexity.

Cluster analysis is a procedure that groups learners on the basis of similarity in patterns of scores. Among the 150 students, four profiles were identified and interpreted on the basis of the clusters' differential performance on the vocabulary and listening tasks, the L1 analytic task, and the L2 analytic and cloze tasks (Table 2). Cluster 1 consisted of those learners who were strongest on all of the measures including those requiring language analysis in both L1 and L2; Cluster 3 were learners who had average scores

*Table 2.* Task performance profile by cluster of learners in Ranta 1998, 2002b

Cluster	L2 Vocabulary L2 and Listening	Analytic Tasks	L1 Analytic Task
1	strong	strong	strong
2	weak	weak	average
3	average	weak	weak
4	weak	weak	weak



on listening and vocabulary tests yet performed poorly on analytic tasks in L1 and L2; Cluster 2 and 4 were below average learners, but the Cluster 2 group was relatively stronger on the L1 metalinguistic task, significantly outperforming the Cluster 3 learners. One explanation of the non-linear relationship between analytic ability and L2 performance (evident in the comparison between the learners in Cluster 2 and 3) is that Cluster 3 represents memory-oriented learners who were able to thrive in the communicative classroom but had difficulty with language analysis.

### **Oral production data**

The participants in this study were drawn from the pool of 50 students who were randomly selected from each of the five September–January ESL classes. In the present analysis, two groups of learners are compared: one group designated as the “analytic” learners and another as the “less analytic” learners. These designations are based on the descriptions of the clusters to which these individuals belonged. That is, the analytic learners are from Cluster 1 and the less analytic learners are from Cluster 2. These two clusters were the only ones available with sufficiently large and comparable cell sizes<sup>6</sup>. There were fifteen learners in each group. The parent clusters to which the analytic and the less analytic groups belonged were significantly different from each other on all of the written L2 measures and on the L1 metalinguistic task. The  $z$  value of each subsample mean was calculated (Welkowitz, Ewan, & Cohen 1991) to confirm that the subsample means were not different from those of the parent cluster at the .05 level.

### **Oral production tasks and analyses**

Two types of oral data were available for these learners: question forms and picture descriptions which provided contexts for third-person possessive determiners. Both of these English structures present challenges to these learners. For example, a student in a pilot study in the same school commented that, at the start of her intensive ESL course, she didn’t know how to say “est-ce que”, and that she would say “sa mother” because she didn’t know the word *her*. More importantly from the perspective of research, both structures are associated with a well-established developmental sequence framework. This means that rather than looking at accuracy in production in terms of being correct vs. incorrect, we have the option of



describing the learner as being at one of six stages for question forms (Pienemann, Brindley, & Johnston 1988) and one of eight for possessive determiners (White 1998). This offers a much greater degree of precision in charting the progress of low level learners. Both of these analytic frameworks have been used in studies conducted with learners in intensive ESL classrooms in Quebec (Spada & Lightbown 1993; White 1998; Lightbown, Spada, & White: this volume).

### Question forms

The descriptions of the developmental stages used in this study (see Table 3) are from Spada and Lightbown (1993), based on the work of Pienemann, Johnston and Brindley (1988). Progression from Stage 1 to Stage 5 is characterized by increasing mastery over the rules for subject-auxiliary inversion in yes/no and wh-questions<sup>7</sup>. Emergence criteria were used rather than accuracy: each question was coded, and the learner's "highest stage" score was determined by identifying the highest stage for which there were at least two different examples.

The Five Questions Picture Card Game was used to elicit question forms (L. White et al. 1991). To play the game, the student is given four similar picture cards and the interviewer chooses one card from a second

*Table 3.* Developmental stages for question formation in English from Spada and Lightbown, 1993

Stage 1	<u>Single words or fragments</u>	A spot on the dog? A ball or a shoe?
Stage 2	<u>SVO with rising intonation</u>	A boy throw the ball? Two children ride a bicycle?
Stage 3	<u>Do-fronting</u>	Do the boy is beside the bus? Do you have three astronaut?
	<u>Wh-fronting</u>	What the boy is throwing? Where the children are standing?
Stage 4	<u>Wh- with copula BE</u>	Where is the ball? Where is the space ship?
	<u>Yes/No questions with aux inversion</u>	Is the boy beside the garbage can? Is there a dog on the bus?
Stage 5	<u>Wh- with auxiliary second</u>	What is the boy throwing? How do you say "lancer" <sup>8</sup> ?

set of the same pictures. The learner is required to ask five questions in order to determine which of the four pictures the interviewer has. There were three different picture sets used in all, not including the first practice set in which the interviewer took the role of the questioner and modeled the type of questions that could be asked. The learner was allowed to ask both yes/no and wh-questions; if the student did not use a wh-question during the first two picture sets, the interviewer prompted for wh-questions (e.g., saying something like, “Ask me a question with *where*”).

### Possessive determiners

The task used to elicit third-person possessive determiners was developed by White (1996) for use with similar ESL students. The learner is asked to describe cartoon pictures. Each cartoon shows a child with one or two parents in the midst of a familiar problem or predicament. For each picture, the interviewer asked the student: “What is the problem?” or “Why is this cartoon funny?” while both student and interviewer looked at the picture. Answering these questions usually led the student to create several obligatory contexts for *his* and *her* (e.g., “She cut her hair and her mother is upset”). Two of the pictures were biased towards contexts for *his* and two were biased towards contexts for *her*.

*Table 4.* Developmental stages in the acquisition of English possessive determiners by francophone learners (Adapted from White 1996: 179)

Stage 1	pre-emergence: avoidance of <i>his</i> and <i>her</i> and/or use of definite article
Stage 2	pre-emergence: use of <i>your</i> for all persons, genders and numbers
Stage 3	emergence of either or both <i>his</i> and <i>her</i>
Stage 4	preference for <i>his</i> or <i>her</i> accompanied by over generalization to contexts for the other form
Stage 5	differentiated use of both <i>his</i> and <i>her</i> not with kin-different gender
Stage 6	agreement rule applied to either <i>his</i> or <i>her</i> kin-different gender
Stage 7	agreement rule applied to both <i>his</i> and <i>her</i> kin-different gender
Stage 8	error-free application of agreement rule to <i>his</i> and <i>her</i> all domains, including body parts

The speech elicited by the second oral production task was analyzed according to the framework of acquisition stages for third person singular possessive determiners developed by White (1996; 1998). The stages in this framework describe the learner's evolving ability to produce *his/her* during a communicative task. The choice between *his* and *her* is difficult for French-speaking learners of English because of the way the two languages differ in their assignment of gender. In French, the choice between masculine and feminine determiners is based on the grammatical gender of the possessed noun (e.g., *Robert voit sa mère*), while in English, the choice depends on the natural gender of the possessor (*Robert sees his mother*). Each student was assigned to a stage according to the descriptions presented in Table 4. The criterion used for determining whether a learner was in stages 4–7 was three instances. In order to facilitate comparisons among learners, each learner was assigned to one of the following categories based on their stage assignment: pre-emergence (stages 1 and 2), emergence (stages 3 and 4), and post-emergence (stages 5, 6, 7) (see White 1998 for rationale for this grouping of stages). For both the question form and the possessive determiner analysis, transcripts of the audio-recordings were coded by a research assistant and then checked by another coder to confirm reliability. All differences in stage assignment were discussed until agreement was reached.

### Fluency analysis

The oral descriptions from one of the cartoons designed to elicit possessive determiners were used for the fluency analysis.<sup>8</sup> This picture depicts a mother's dismay at her daughter's self-administered haircut. Fluency was measured using quantitative measures of temporal aspects of fluency such as speech rate, pausing behaviour and dysfluency markers. These well-established techniques for measuring fluency (e.g., Lennon 1990; Raupach 1980; Towell, Hawkins, & Bazergui 1996; Riggenbach 1991) provide a greater degree of precision than is possible with a rating scale such as the one developed by ACTFL for elementary school students (Swender & Duncan 1998) and therefore maximizes the likelihood of detecting differences among learners.

Each student's picture description was isolated from the audiotape and the computer program *SoundEdit* was used to digitize the audiorecordings and to calculate the students' speech rates, pause times, and "mean length of run". The duration of the speech segments and pauses were measured

separately. Filled pauses such as “um” or “uh” and unfilled pauses were also measured separately but a combined score was used for the Pauses variable used in the analysis. Speech rates were calculated and recorded as the number of syllables per second; pauses and filled pauses were timed and averaged for each student; and mean length of run, defined as the amount of speech between pauses, was calculated by measuring the average length of speech in seconds between pauses of a duration of 25 milliseconds or longer.

The transcriptions of the learners’ cartoon descriptions were also coded for the following dysfluency markers:

- false start: an incomplete utterance or a change in structural organization.  
Example: *it’s not uh the mother’s not happy*
- reformulation: lexical or grammatical self-correction.  
Examples: *in his left arm uh hand; the her mother*
- repetition: exact repetition of speaker’s own immediately preceding form.  
Example: *the the mother*

All of the transcripts of the learners’ descriptions were coded by a research assistant; the author verified the coding of a sub-set of the first ten transcripts. Differences in the application of the coding categories were discussed until consensus was reached. When problems or questions arose during the coding process, the problems were tagged and discussed by both researchers to determine an optimal coding. The frequency of the dysfluency markers was tallied for each student and then standardized by dividing the frequency counts by the sum of the total number of syllables and filled pauses. T-tests were performed to determine significant differences between the two groups. The alpha was set at .05.

## RESULTS

### Interlanguage analysis

Figure 1 displays the results for the question stage analysis. Note that, as there was only one stage 5 score (in the less analytic group), stages 4 and 5 have been combined. In the analytic group, no learners were in stage 2 and a large majority of the learners produced Stage 4 questions (12 or 80%). In contrast, a few of the less analytic learners (2 or 13%) have SVO as their highest stage (stage 2) and a smaller proportion of the learners are in the

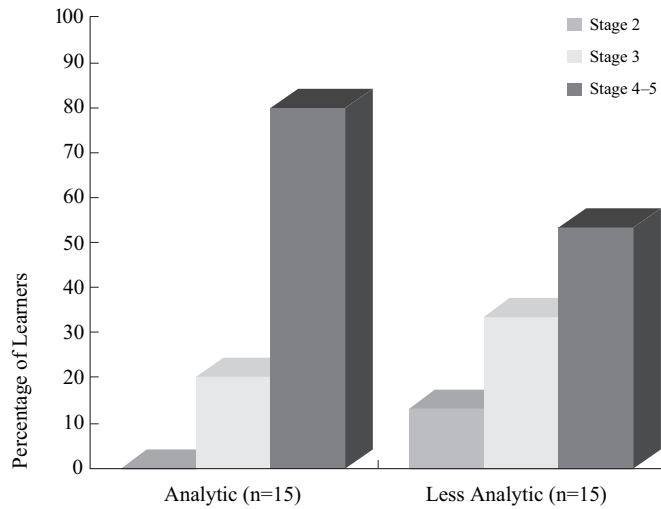


Figure 1.

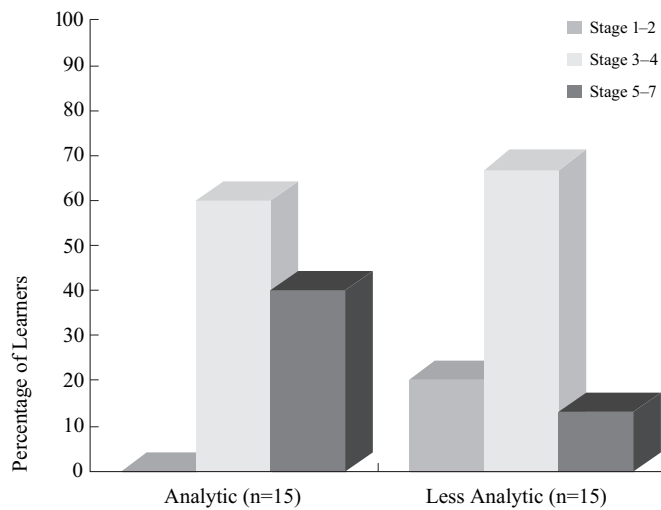


Figure 2.

highest stages (53%). This difference in distribution is not significant according to the Mann-Whitney U test of significance ( $z = -.902, p = .461$ ).

Figure 2 displays the results for the possessive determiner stages. A similar percentage of learners in the analytic and less analytic groups are in the Emergence stages (60% and 67%, respectively). But the analytic group has a larger percentage of learners in the higher stages; 6 (40%) are in the

Post-emergence stages whereas this is the case for only 2 (13%) of the less analytic learners. No learners in the analytic group are in the Pre-emergence stages but 3 (20%) of the less analytic learners are. The difference between groups here is statistically significant according to the Mann-Whitney U test of significance ( $z = -2.324$ ,  $p = .033$ ).

### Fluency analyses

Table 5 presents the descriptive statistics for the results from the six measures of fluency performed on the speech samples. A fluent speaker will be expected to have a high score for syllables/second and mean length of run variables but a low score for the false start, reformulation, repetition, and pausing categories. Although the means for the analytic group generally indicate greater fluency for that group, there was only one significant difference: the less analytic learners used significantly more repetitions than the analytic learners ( $p = .01$ )

*Table 5.* Descriptive statistics for fluency measures (standard deviations in parantheses)

	syllables/ second	fills and pauses	mean length of run	false starts	reformu- lations	repeti- tions
Analytic Group*	2.17 (.55)	3.29 (2.99)	1.41 (.25)	.39 (.68)	1.89 (1.71)	1.43# (1.48)
Less analytic*	2.12 (.54)	3.25 (1.88)	1.29# (.27)	.97 (.82)	2.11 (1.73)	4.09# (3.19)

\* n=15, # indicates n=14

The relationship among the various measures of fluency is complex, as previous fluency research has shown (Riggenbach 1991). In order to get insight into the learners' performance, the fluency scores from two learners from each group were examined more closely; Learners A and B are from the analytic group, C and D from the less analytic group. These learners were chosen to illustrate the range within the groups (and not the extremes) in terms of accuracy and metalinguistic ability. Learner A and B were in stages 7 and 6 respectively for possessive determiners and stage 4 for questions. Learner C was at stage 5 on possessive determiners and stage 4 for

questions, whereas Learner D was at stage 4 for possessive determiners and stage 3 for questions.

Table 6. Individual patterns on fluency measures (ranks in parentheses)

	syllables/ second	fills and pauses	mean length of run	false starts	reformu- lations	repeti- tions
Learner A	1.84 (2)	2.44 (1)	1.66 (2)	0 (1)	4.72 (4)	0 (1)
Learner B	1.44 (4)	4.63 (4)	1.41 (4)	0 (1)	2.78 (2)	1.39 (2)
Learner C	2.25 (1)	2.83 (2)	1.87 (1)	.64 (3)	.64 (1)	7.01 (3)
Learner D	1.46 (3)	3.37 (3)	1.45 (3)	1.54 (4)	3.08 (3)	7.69 (4)

Table 6 presents the results for the four selected learners. In the table, the rankings among these four learners on each measure are indicated in parentheses. Note that a high ranking on the dysfluency measures means having low rates of occurrence. The pattern among these four learners illustrates the complexity of fluency as a concept. For example, Learner C, one of the less analytic learners has the largest number of first place rankings relative to the other three. She speaks quickly but also has a high rate of repetitions. Learner A, an analytic learner, ranks first or second on all but the reformulation category. Examination of the transcript (Appendix A) reveals that these reformulations are grammatical in nature and thus provide further evidence of an accuracy orientation in this learner. Learner B's pattern of scores looks inconsistent; he ranks well with respect to repetitions, reformulations and false starts, but low with respect to pausing and to the speech rate variables. Examination of the transcript suggests that concern for accuracy might be responsible for his performance. He is highly accurate and shows evidence of greater grammatical complexity than is common among learners in intensive programs. For example, Learner A uses the progressive appropriately, demonstrates knowledge of subject-verb agreement (*is he showing vs. are they talking*), and can invert both noun phrases and pronouns as subjects. There is also an absence of inappropriate do-fronting. From the transcript it seems as if he is checking the response of

the interviewer before each utterance. This leads to a high rate of pausing at sentence boundaries, which is misleading in that it is not an indication of dysfluency. Learner D's description is also less monologue-like than those of Learners A and C which may influence the rate of pausing. However, he ranks low across all the measures and he is the least able to communicate effectively what is going on in the picture.

The fine-grained analysis of the fluency measures from these four learners leads to textbook examples of well-established learner profiles (Skehan 1998): Learner A is both accurate and fluent; Learner B is the analytic learner who sacrifices fluency for accuracy; Learner C looks like a fluency-oriented learner who sacrifices accuracy for fluency, and Learner D is low in fluency because of low proficiency in the L2. Whether these are enduring profiles of these learners cannot be determined on the basis of the available data. However, this microanalysis serves to shed light on the lack of difference in the fluency results since the high and low fluency individuals cancel each other out in both groups.

## DISCUSSION

The results from the analysis of the oral production of more versus less analytic learners in an intensive ESL program reveal that language analytic ability is associated with interlanguage development but not with measures of fluency. The results for the possessive determiners patterned such that learners in the analytic group were more likely to be in the higher stages of development and less likely to be in the lower stages. Since the instructional content which all of the teachers followed closely did not include form-focussing activities and grammatical explanations, we must conclude that these learners acquired this grammatical rule on the basis of the processes of input-processing, along the lines described in Skehan (2002).

The results from the question form analysis, while following the same trend, did not lead to a statistically significant difference between the two groups. One explanation for this is that learners often use chunks when asking questions which results in less reliable stage assignments. For example, a learner may produce stage 4 questions such as "Where is X?" or "Is there X?" without being able to form other questions with the copula. Thus the stage 4 assignment may overestimate the learner's knowledge of the rules for inversion in questions in English. This is a well known problem when coding question data using emergence criteria (Spada, Lightbown & White, this volume). In many ways, the pattern of yes/no questions are



more informative than the wh-questions on this type of task (Ranta 2002). For example, if we examine the yes/no questions from Learners A, B, C, and D (Table 7), we see that the highest stage assignment masks differences in the complexity of the forms of learners' questions. Among these four learners, we see a greater degree of hypothesis-testing in the questions produced by the analytic learners (A and B) than by the less analytic learners (C and D).

Table 7. Yes/no questions produced by four learners

Learner A Analytic group	Learner B Analytic group	Learner C Less Analytic group	Learner D Less Analytic group
is he showing something?	is there a garbage?	have a garbage can?	the boy wearing?
is the dog run for a ball?	is it on the right corner?	there is a ballon on the picture?	the dog have it a spot?
is there a garbage?	is she on the sidewalk?	if there is a shoes?	do you have...
is the dog have dots?	is she on the grass?	there is a boy with a sweater?	did you have...
is there two people or one?	do the dog have...	there is a girl and a boy who do a bicycle?	he are in the corner or at the front?
is there a cat?	do the dog is...	there is a dog?	the sun are the left or at the right?
is the bus driver is a woman?	do the bicycle rider is holding his bike?	is this on the sidewalk?	do you have...
is it on in the middle?	do the two woman are talking together?	do you have...	the two woman talk is in the road?
is there only one person?		if there are many traffic?	you have some people go out?
is three some people who look from the building?			you have some people beside the woman?
are they talking to each other?			

The interlanguage analyses, particularly the possessive determiners, provide support for the notion that analytic learners acquire grammatical rules more rapidly from the input. This is in keeping with the definition of aptitude as “speed of learning” (Carroll 1962, 1981) and with the specific hypotheses generated by Skehan (2002). This finding parallels the conclusion drawn by Lightbown and Spada (1997; Spada, Lightbown, & White, this volume) that form-focussed instruction causes more learners to leave lower developmental stages and attain higher ones more quickly.<sup>9</sup> In other words, the “internally generated saliency” (Sharwood Smith 1991) of the analytic learners in this study appears to have had the same impact on learning as

the “externally generated saliency” provided by form-focussed instruction. This similarity in outcome substantiates the description of the noticing stage as proposed by Skehan (2002: 88).

The second research question examined whether analytic ability offers any advantage to learners in their development of fluency. To address this question a number of measures of fluency were used, since no single indicator can provide a complete picture. The two groups were not significantly different from each other in terms of speech rate, mean length of run, mean length of pauses, and frequency of reformulations and false starts. The only significant difference was in the frequency of repetitions. The analytic group produced significantly fewer immediate repetitions. This is a dysfluency marker that is identified by Dörnyei and Kormos (1997) as a problem-solving mechanism related to time pressure. According to their framework, repetition is a stalling tactic that is used when learners feel that their processing speed is not up to the demands of the situation. This suggests that encoding processes of the less analytic group were less automatized than those of the analytic group. However, the speech rate and mean length of run variables, both of which are believed to reflect automaticity in a fairly direct way (Schmidt 1992: 362; Towell et al. 1996), did not reveal any statistically significant differences between the groups. It is quite likely, then, that this result is due to chance rather than a reflection of real group differences in fluency. A longer speech sample from a larger number of learners is clearly needed to clarify this point.

The relatively small range in scores and low standard deviations for the mean length of run in both the analytic and the less analytic groups supports the notion that temporal aspects of fluency reflect the learning environment and not individual differences in analytic ability or overall L2 proficiency. The fact that there were no significant differences between the groups, despite having significantly different levels of proficiency in English with respect to written proficiency measures, suggests that the kind of instruction the learners received in the intensive program was transfer-appropriate for fluency (Segalowitz 2000). Students in the intensive program engaged in small group activities that allowed them to practice vocabulary and formulaic expressions. Since all students were given ample opportunities to participate in a wide variety of oral activities such as games, skits, and presentations, they had similar opportunities for proceduralization to take place.

One strength of this study is that the two learner groups were significantly different from each other in terms of performance on analytic tasks, but are similar in terms of L1, age, and L2 learning background. In other widely reported aptitude studies, analytic ability was confounded with

other variables such as age in Skehan (1986), age at onset of L2 learning plus type of instructional approach in Harley and Hart (1997), and age, gender and level of education in Wesche (1981).

There are, however, important limitations to this study. The speech samples were very brief (ranging from a low of 21 syllables to a high of 163). The learner groups were also relatively small, so the findings may not generalize to the wider population of ESL learners in intensive programs. The work of Skehan and Foster (1999) reminds us that fluency is not a learner attribute but a reflection of their knowledge under certain conditions. From the results of other studies, it is possible that the narrative task used here may be biased against eliciting fluent speech (Ejzenberg 2000; Skehan & Foster 1999); a main task effect might have washed out fluency differences among the learners. Lastly, although L2 teachers are likely to be gladdened by this demonstration that communicative teaching provides equal opportunities for the development of fluency, it may well be the case that qualitative ratings would offer a different picture of the abilities of individual students. Thus, we can conclude that reports of the demise of aptitude in SLA are exaggerated, but there is a need for research to expand and refine existing conceptions of aptitude in the light of what we know about the acquisition process. Rather than being definitive, the results of this study should be considered an invitation to further investigation into the effect of learner variables and of instructional features on both fluency and interlanguage development in a L2.

### **Acknowledgements**

This paper is based on my dissertation study which was funded by grants from Concordia University and FCAR to me and by grants from FCAR and SSHRC to Patsy Lightbown and Nina Spada. It also reflects what I have learned after completion of the dissertation from my rewarding collaborations with Joanna White and with Tracey Derwing. The fluency analyses would not have been possible without Tracey's guidance and the careful attention to detail of Marilyn Abbott at the University of Alberta. I also benefited greatly from feedback on the original draft of this paper from Roy Lyster and Joanna White.

## Appendix

Transcripts for Learners A,B, C and D for oral description of “Haircut”  
Picture Conventions:

+ learner speech; = researcher’s speech; [ ] French words; //

### *Learner A*

+ The mother’s crying and uh her little baby is asking her what’s the matter with her and uh it’s because the mother saw that the little baby cut her hair, and uh there’s lots of hair on the floor. She’s got scissors in their hand, and the little baby is surprised to see her mother crying. And the her mother is on the floor, sitting on the floor on her knees and uh she have the hands \*on in\* \*their in her hair\*, uh, she’s in socks she don’t have any uh shoes, and uh her little baby’s have \*the her\* hands on her mother’s arm.

### *Learner B*

+ The mother is crying. The little girl is holding a scissor in his left arm, uh hand. The mother’s hand is on her ears.  
= Mmhm.  
+ The mother is sitting on her knees.  
= Mmhm.  
+ The little girl is talking to his mother.  
= Mmhm. What’s on the floor?  
+ On the floor, there’s a lot of hairs, the little girl’s hair.

### *Learner C*

+ Uh the the the ma- the g- the girl the the mother of the the girl cry because uh the little girl cut her hair and uh the little girl said “What’s the matter Mama? Don’t you like my haircut?” and the mother I don’t think laughs she’s very happy of that um, the the woman are /kinees/ how do you say [à genouiller]? [A genouiller]?  
= Uh kneeling.  
+ Knee- kneeling, um and the girl have a scissor \*on his hand on her hand\*. The the mother are are surprised too I think, the girl have a little uh, fuzzy not fuzzy, I don’t know a little haircut you see the hair uh cut is a bubble for the the sentence of the little girl.

*Learner D*

- + [Ben] the mother cry.
- = Right, very good. What happened? What did she – why is she crying?
- + Maybe she, because they have a cut hair and sh- she uh she, [ben] the person cut \*his haircut his hair\* uh don't know. Don't make it /jeune? / and I don't know. He he make a hurt there her head uh maybe.
- = I'm sorry?
- + He make a hurt uh his hair, %maybe%.
- = %Make% make hard?
- + Make hard.
- = Yeah it's not uh, the mother's not happy is she?
- + No.

**Notes**

1. I am inspired by the definition of metalinguistic development of Ryan and Ledger (1984: 157) who characterize it as: “the gradual shift of attention from meaning to structure in tasks requiring deliberate control over language forms. The essential feature of the beginnings of linguistic awareness seems to be flexibility of strategy – the ability to decenter, to shift one's focus from the most salient attribute of a message (its meaning and contextual setting) to structure (the ordinarily transparent vehicle by which meaning is conveyed).”
2. Reves (1983) found aptitude including analytic ability to be a predictor of performance in oral production in Hebrew as a L2 among Arabic-speaking high school students living in a bilingual community in Israel. Although this study is often cited as proof that aptitude influences L2 learning in *informal* settings (e.g., Skehan 1989, 2002), such a conclusion overlooks the fact that these learners had had formal instruction in Hebrew which focussed on accuracy in the written language from grade 5 onwards.
3. Recent studies of extraversion (Dewaele & Furnham 1999, 2000) and field independence/dependence (Johnson, Prior, & Artuso 2000) suggest that personality traits influence cognitive processes which have an impact on L2 fluency.
4. The many versions of the model have slightly different designations (i.e., ACT\*, ACT-R) but they do not differ greatly on the points discussed here (see Anderson 1993 for comparisons of old and new models).
5. This task required learners to read a coherent text and were asked to cross out any errors that they noticed. The text contained 32 errors of which 21 were possessive determiners and 11 involved different aspects of English. The metalinguistic task cannot be performed accurately without having an analyzed representation (Bialystok 1994) of the English gender agreement rule (see White & Ranta 2002 for discussion of Bialystok's theories with respect to this task

and to oral production of *his/her*). In the context of communicative instruction without focus on form, it is assumed that those who have language analytic ability will be able to carry out this process of analysis more rapidly than less analytic peers.

6. The random selection of learners took place in January at the end of the ESL program for the Fall cohort. The cluster analysis was carried out after the final data collection which took place in June. The cluster distribution of learners selected for the oral tasks were as follows: 15 in Cluster 1, 18 in Cluster 2, 8 in Cluster 3, and 5 in Cluster 4. The data from three learners in Cluster 2 were removed (for the analysis presented here) in order to create balanced groups: one was removed because of poor sound quality; one was a clear outlier in terms of his interaction with the interviewer; and one was selected for deletion using a table of random numbers.
7. The task did not offer opportunities for indirect questions which are found in Stage 6 of Pienemann et al's (1988) framework.
8. The analysis of the fluency measures for the whole sample of 97 students who participated in the larger study (Lightbown & Spada 1997) are presented in Ranta, Abbott, & Derwing (in preparation).
9. In fact, they compared the results from the 100 students at the Intensive ESL school with the results from previous experimental studies on question forms (White, Spada, Lightbown, & Ranta 1991; Spada & Lightbown 1993) and on possessive determiners (White 1998).

## References

- Anderson, J. R.  
 1983 *The architecture of cognition*. Cambridge, MA: Harvard University Press.  
 1993 *Rules of the mind*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Bialystok, E.  
 1994 Analysis and control in the development of second language proficiency. *Studies in Second Language Acquisition* 16: 157–168.
- Brumfit, C.  
 1984 *Communicative methodology in language teaching*. Cambridge: Cambridge University Press.
- Carroll, J. B.  
 1962 The prediction of success in intensive foreign language training. In *Training research and education*, R. Glaser (ed.), 87–136. New York: John Wiley.  
 1981 Twenty-five years of research on foreign language aptitude. In *Universals in language learning aptitude*, K. Diller (ed.), 83–118. Rowley, MA: Newbury House.

- Carroll, J. B. and S. Sapon  
1958 *The Modern Language Aptitude Test – Form A*. NY: The Psychological Corporation.
- Chapelle, C. and P. Green  
1992 Field independence/dependence in second-language acquisition research. *Language Learning* 42: 47–83.
- Cook, V.  
1996 Reprint. *Second language learning and language teaching*. 2d ed. London: Edward Arnold.
- DeKeyser, R.  
1997 Beyond explicit rule learning: Automatizing second language morpho-syntax. *Studies in Second Language Acquisition* 19: 195–221.  
1998 Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar. In *Focus on form in classroom second language acquisition*. C. Doughty, and J. Williams (eds.), 42–63. Cambridge: Cambridge University Press.  
2002 Automaticity and automatization. In *Cognition and second language instruction*. P. Robinson (ed.), 125–151. Cambridge: Cambridge University Press.
- Dewaele, J. M. and A. Furnham  
1999 Extraversion: The unloved variable in applied linguistic research. *Language Learning* 49: 509–544.
- Dörnyei, Z. and J. Kormos.  
1998 Problem-solving mechanisms in L2 communication: A psycholinguistic perspective. *Studies in Second Language Acquisition*, 20, 349–385.
- Ehrman, M. and R. Oxford  
1995 Cognition plus: Correlates of language learning success. *Modern Language Journal*, 79, 67–89.
- Ejzenberg, R.  
2000 The juggling act of oral fluency: A psycho-sociolinguistic metaphor. In *Perspectives on fluency*, H. Riggenbach (ed.), 287–313. Ann Arbor: University of Michigan Press.
- Evans, M. A.  
1985 Self-initiated speech repairs: A reflection of communicative monitoring in young children. *Developmental Psychology* 21: 365–377.
- Foster, P. and P. Skehan  
1996 The influence of planning on performance in task-based learning. *Studies in Second Language Acquisition* 18: 299–324.
- Freed, B.  
1995 What makes us think that students who study abroad become fluent? In *Second language acquisition in a study abroad context*, B. Freed (ed.), 123–148. Amsterdam/Philadelphia: John Benjamins.

- Gass, S.  
 1997 *Input, interaction, and the second language learner*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Gardner, R. and W. Lambert.  
 1972 *Attitudes and motivation in second-language learning*. Rowley, MA: Newbury House.
- Gatbonton, E. and N. Segalowitz  
 1988 Creative automatization: Principles for promoting fluency within a communicative framework. *TESOL Quarterly* 22: 473–491.
- Hadley, A.O.  
 1993 *Teaching language in context*. 2d (ed.) Boston, MA: Heinle and Heinle.  
 2001 *Teaching language in context*. 3d (ed.) Boston, MA: Heinle and Heinle.
- Harley, B. and Hart, D.  
 1997 Language aptitude and second language proficiency in classroom learners of different starting ages. *Studies in Second Language Acquisition* 19: 379–400.
- Harley, B. and D. Hart  
 2002 Age, aptitude, and second language learning on a bilingual exchange. In *Individual differences and instructed language learning*, P. Robinson (ed.), 301–330. Amsterdam/Philadelphia: John Benjamins.
- Horwitz, E.  
 1987 Linguistic and communicative competence: Reassessing foreign language aptitude. In *Foreign language learning*, B. VanPatten, T. Dvorak, and J. Lee (eds.), 146–157. Cambridge, MA: Newbury House.
- Johnson, J., Prior, S. and M. Artuso  
 2000 Field dependence as a factor in second language communicative production. *Language Learning* 50: 529–567.
- Koponen, M. and H. Riggenbach.  
 2000 Overview: Varying perspectives on fluency. In *Perspectives on fluency*, H. Riggenbach (ed.), 5–24. Ann Arbor: University of Michigan Press.
- Kormos, J.  
 1999 Monitoring and self-repair. *Language Learning* 49: 303–342.  
 1999 The effect of speaker variables on the self-correction behaviour of L2 learners. *System* 27: 207–221.
- Krashen, S.  
 1981 Aptitude and attitude in relation to second language acquisition and learning. In *Individual differences and universals in language learning aptitude*, K. Diller (ed.), 155–175. Rowley, MA: Newbury House.
- Lennon, P.  
 1990 Investigating fluency in EFL: A quantitative approach. *Language Learning* 40: 387–417.



- 2000 The lexical element in spoken second language fluency. In *Perspectives on fluency*, H. Riggenbach (ed.), 25–42. Ann Arbor: University of Michigan Press.
- Lightbown, P.M. and N. Spada
- 1990 Focus-on-form and corrective feedback in communicative language teaching. *Studies in Second Language Acquisition* 12: 429–448.
- 1994 An innovative program for primary ESL in Quebec. *TESOL Quarterly* 28: 563–579.
- 1997 Learning English as a second language in a special school in Québec. *Canadian Modern Language Review* 53: 315–355.
- McLaughlin, B. and R. Heredia
- 1996 Information-processing approaches to research on second language acquisition and use. In *Handbook of second language acquisition*, W. Ritchie, and T. Bhatia (eds.), 213–228. NY: Academic Press.
- Obler, L.
- 1989 Exceptional second language learners. In *Variation in second language acquisition vol. 2: Psycholinguistic issues*, S. Gass, C. Madden, D. Preston, and L. Selinker (eds.), 141–159. Clevedon/Philadelphia: Multilingual Matters.
- Pienemann, M., M. Johnston and G. Brindley
- 1988 Constructing an acquisition-based procedure for assessing second language acquisition. *Studies in Second Language Acquisition* 15: 495–503.
- Pimsleur, P.
- 1966 *The Pimsleur language aptitude battery*. NY: Harcourt Brace Jovanovich.
- Ranta, L.
- 1998 *Focus on form from the inside: The significance of grammatical sensitivity for L2 learning in communicative ESL classrooms*. Ph.D. diss., Concordia University, Montreal.
- 2002a *The concept of chunking in L2 acquisition*. Paper presented at the annual meeting of the American Association of Applied Linguistics, Salt Lake City, UT.
- 2002b The role of learners' language analytic ability in the communicative classroom. *Individual differences and instructed language learning*, In P. Robinson (ed.), 159–80. Amsterdam/Philadelphia: John Benjamins.
- Ranta, L., M. Abbott and T. Derwing
- in prep Exploring fluency development in early second language learning.
- Raupach, M.
- 1980 Temporal variables in first and second language speech production. In *Temporal variables in speech: Studies in honour of Frieda Goldman-Eisler*, H. Dechert, and M. Raupach (eds.), 263–270. The Hague: Mouton.

- Reves, T.  
1983 *What makes a good language learner?* Ph.D. diss., Hebrew University of Jerusalem.
- Riggenbach, H.  
1991 Towards an understanding of fluency: A microanalysis of nonnative speaker conversations. *Discourse Processes* 14: 423–441.
- Ryan, E. B. and G. Ledger  
1984 Learning to attend to sentence structure: Links between metalinguistic development and reading. *Language awareness and learning to read*, In J. Downing, and R. Valtin (eds.), 151–171. New York: Springer-Verlag.
- Sawyer, M. and L. Ranta  
2001 Aptitude, individual differences and L2 instruction. *Cognition and second language instruction*, In P. Robinson (ed.), 319–353. Cambridge University Press.
- Scarcella, R. and R. Oxford  
1992 *The tapestry of language learning: The individual in the communicative classroom*. Boston, MA: Heinle and Heinle.
- Schmidt, R.  
1992 Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition* 14: 357–385.
- Segalowitz, N.  
1997 Individual differences in second language acquisition. In A. M. B. de Groot, and J. F. Kroll (eds.), *Tutorials in bilingualism: Psycholinguistic perspectives*, 85–112. Hillsdale, NJ: Lawrence Erlbaum.  
2000 Automaticity and attentional skill in fluent performance. In *Perspectives on fluency*, H. Riggenbach (ed.), 200–219. Ann Arbor: University of Michigan Press.
- Sharwood Smith, M.  
1991 Speaking to many minds. *Second Language Research* 7: 118–132.
- Shrum and Glisan  
2000 *The teacher's handbook*. 2d ed. Boston, MA: Heinle and Heinle.
- Skehan, P.  
1986 Cluster analysis and the identification of learner types. *Experimental approaches to second language acquisition*, In V. Cook (ed.), 81–94. Oxford: Pergamon.  
1989 *Individual differences in second-language learning*. London: Edward Arnold.  
1990 The relationship between native and foreign language learning ability: Educational and linguistic factors. *Current trends in European second language acquisition research*, In H. Dechert (ed.), 83–106. Clevedon, UK: Multilingual Matters.  
1991 Individual differences in second language learning. *Studies in Second Language Acquisition* 13: 275–298.

- 1998 *A cognitive approach to language learning*. Oxford: Oxford University Press.
- 2002 Theorising and updating aptitude. In *Individual differences and instructed language learning*, P. Robinson (ed.), 69–93. Amsterdam/Philadelphia: John Benjamins.
- Skehan, P. and P. Foster
- 1999 The influence of task structure and processing conditions on narrative retellings. *Language Learning* 49: 93–12.
- Spada, N.
- 1990 Observing classroom behaviours and learning outcomes in different second language programs. In *Second language teacher education*, J. C. Richards, and D. Nunan (eds.), 293–310. Cambridge: Cambridge University Press.
- Spada, N. and M. Fröhlich
- 1995 *The Communicative Orientation of Language Teaching COLT observation scheme: Applications for research and teacher education*. Sydney: National Centre for English Language Teaching and Research NCELTR.
- Spada, N. and P. M. Lightbown
- 1989 Intensive ESL programs in Quebec primary schools. *TESL Canada Journal* 7: 11–32.
- 1993 Instruction and the development of questions in L2 classrooms. *Studies in Second Language Acquisition* 15: 205–224.
- Swender, E. and G. Duncan
- 1998 ACTFL performance guidelines for K-12 learners. *Foreign Language Annals* 31: 479–491.
- Towell, R., Hawking, R. and N. Bazergui
- 1996 The development of fluency in advanced learners of French. *Applied Linguistics* 17: 84–119.
- Tucker, R., Hamayan, E. and F. Genesee
- 1976 Affective, cognitive and social variables in second language acquisition. *Canadian Modern Language Review*, 23: 214–226.
- Verhoeven, L.
- 1989 Monitoring in children's second language speech. *Second Language Research* 5: 141–155.
- Weary, K.
- 1987 *An evaluation of the pedagogical materials of an intensive ESL program used by a francophone Quebec schoolboard at the grade 6 level*. Unpublished M. Ed. monograph, McGill University, Montreal.
- Welkowitz, J., Ewen, R. and J. Cohen
- 1991 *Introductory statistics for the behavioral sciences*. San Diego: Harcourt Brace Jovanovitch.
- Wells, G.
- 1985 *Language development in the pre-school years*. Cambridge: Cambridge University Press.

- Wesche, M.  
1981 Language aptitude measures in streaming, matching students with methods, and diagnosis of learning problems. In *Individual differences and universals in language learning aptitude*, K. Diller (ed.), 119–154. Rowley, MA: Newbury House.
- White, J.  
1996 *An input enhancement study with ESL children: Effects on the acquisition of possessive determiners*. Ph.D. diss., McGill University, Montreal.  
1998 Getting learners' attention: A typographical input enhancement study. In *Focus on form in classroom second language acquisition*, C. Doughty, and J. Williams (eds.), 85–113. Cambridge: Cambridge University Press.
- White, J. and L. Ranta  
2002 Examining the interface between metalinguistic task performance and oral production in a second language. *Language Awareness* 11 (4): 259–290.
- White, L., N. Spada, P. M. Lightbown and L. Ranta  
1991 Input enhancement and L2 question formation. *Applied Linguistics* 12: 416–432.
- Wong Fillmore, L.  
1979 Individual differences in second language acquisition. In *Individual differences in language ability and language behavior*, C. Fillmore, D. Kempler, and W. Wang (eds.), 203–228. NY: Academic Press.

## **Formal instruction and the acquisition of verbal morphology**

*Kira Gor and Tatiana Chernigovskaya*

The present study investigates the processing of complex verbal morphology in second language (L2) learners. It focuses on the role of input frequencies, morphological complexity, and morphological cues in L2 acquisition of inflectional morphology in a formal instructional setting, and compares the L2 processing data to the baseline native language (L1) data. In particular, the study addresses the following research issues: (1) Does explicit instruction in complex morphological rules result in the successful learning of these rules as reflected on pencil-and-paper tests? (2) Does explicit instruction on verb conjugation facilitate the development of native-like verbal processing strategies in L2 learners (3) What is the role of input frequencies in L2 processing of complex verbal morphology? Two groups of subjects, adult American formal learners and adult native speakers of Russian, participated in the experiments, which involved oral and written generation of the non-past tense Russian novel verb forms from the past-tense stimuli. The study uses its own L2 input frequency counts obtained for the L2 participants. The results of the study indicate that both groups of subjects used the same processing strategies, and relied on default processing in similar ways. The differences in the rates of L2 and L1 use of the individual conjugational patterns are to a considerable extent due to the differences in the input frequencies to the L2 and L1 speakers. Neither the dual-system nor the single-system theories of morphological processing can handle the reported data on the processing of complex Russian morphology, which calls for a model integrating both of these theories. The influence of input frequencies on L2 verbal processing documented in this study highlights the importance of the statistical characteristics of the language used in a formal classroom for the development of native-like processing strategies in L2 learners.

### 1. Focus on form in formal instruction

In the last decade, SLA research has experimentally tested Stephen Krashen's hypothesis that comprehensible input is not only necessary, but sufficient for SLA (Krashen 1977). Research demonstrated that while comprehensible input itself is insufficient for learning to take place (Trahey and White 1993), some properties of the input and of L2 processing contribute to input becoming intake. For example, the learner needs to be aware of the structural features of the input which is to be acquired; as according to Richard Schmidt, without noticing there can be no learning (Schmidt 1995a)<sup>1</sup>. A distinction between **focus on forms** adopted in traditional methodologies when forms are taught in isolation as grammatical paradigms, and **focus on form**, attention to linguistic form within a meaningful context, which was drawn by Michael Long, influenced the research agenda (Long 1991; Long and Robinson 1998)<sup>2</sup>. Numerous data indicate that in a meaning-focused classroom, focus on the formal properties of language, whether one uses explicit explanations (Spada and Lightbown 1993; Alanen 1995; Robinson 1995, 1996) or specifically structured input, drawing the learner's attention to the formal properties of language (VanPatten and Cadierno 1993; VanPatten and Oikkenon 1996; VanPatten 1996), contributes to successful acquisition of linguistic structures.

Michael Sharwood Smith suggested that one possible way to raise language learners' consciousness is to use enhanced input to the learner, e.g., typographical modifications of target language structures in a written text (Sharwood Smith 1993). In fact, such typographically enhanced input was shown to have a positive influence on learners' output in a second semester Spanish class (Jourdenais et al. 1995). In another study, input enhancement increased student accuracy in Spanish within a content-based instructional setting. The enhanced version of the reading passage administered to the Focus on Form group contained all the preterit and imperfect verbs underlined and colour-coded. Within this setting, the Focus on Form group outperformed the group, which received purely communicative instruction (Leeman et al. 1995). The positive effect of focus on form within a communicative classroom was also found in a Canadian study of the acquisition of several English structures by 10–12-year-old francophone students in an intensive ESL course (Spada and Lightbown 1993). Study of a similar student sample in a similar learning environment revealed that flooding learners for 2 weeks with specially prepared materials on adverb placement in English without any form-focused instruction was not sufficient to drive out forms that are permitted in French but not in English from the students'

output. This conclusion is based on students' performance on a battery of tests ranging from grammaticality judgment to oral production tasks (Trahey and White 1993). Interestingly enough, even young children, whose cognitive abilities and metalinguistic awareness are much more limited than adults', benefited from focus on form within a communicative, content-oriented French immersion classroom (Harley 1993, 1998).

Although instruction, which provides explicit explanations of grammar rules, especially simple rules involving transparent form-function relations, proves beneficial to adult learners (Alanen 1995; Robinson 1995; Williams and Evans 1998), the positive role of implicit instruction remains to be proven empirically. In fact, research has failed to show the positive influence of implicit instruction so far (Ellis 1993; White 1998). For example, extensive amounts of implicit training (with no explanations) had no positive effect on the acquisition of the rule for soft mutation in Welsh (Ellis 1993). As for explicit instruction, several studies have experimentally demonstrated its advantages. Riika Alanen studied the acquisition of Finnish locative suffixes by 4 groups of beginning learners who were native speakers of English. On most tests she obtained a clear advantage for the rule only and rule plus enhanced input groups over the control and enhanced input only groups (Alanen 1995). Peter Robinson reported similar results for the acquisition of easy rules in ESL: the explicitly instructed group outperformed three other non-instructed groups (Robinson 1995). At the same time, the literature suggests that grammar explanations provided to the learner without any practice may result in relatively poor performance (VanPatten and Cadierno 1993; VanPatten and Oikkenon 1996). Therefore, activities promoting language use are a necessary condition for explicit input to become intake.

To summarize the above, recent literature on instructed SLA provides numerous examples of the positive influence that focus on form, or **FonF** instruction (explicit explanations and enhanced or structured input) has on learners' performance (Doughty and Williams 1998)<sup>3</sup>. However, there are no experimental data on the role of formal instruction in developing native-like processing strategies, which go beyond the direct application of the rules explicitly taught and practiced in a language classroom. Our study is devoted to the structure of the mental lexicon and processing strategies in American learners of Russian shaped exclusively by the classroom experience. It focuses on the following issue: What is accomplished in a beginning language classroom, not in terms of committing to memory a set of target language forms and patterns, but in terms of developing native-like intuitions about how the target language system works?



Two aspects of the beginning Russian language classroom shaped the learners taking part in this study. First and foremost, it is a communicative classroom. The course our subjects took is video-based, and the soap-opera-like video with its characters and story line provides motivation for extensive conversational practice. And second, it delivers focus on form instruction with explicit explanations and practice of language rules.

## **2. The processing of verbal morphology**

The general framework for this study comes from research on the processing of English past-tense regular and irregular verbs, which raised the issue of modularity in morphological processing. The modular, or dual-system approach, claims that regular and irregular verbs are processed by two distinct mechanisms. Regular verb forms are computed in a rule-processing system, while irregular verbs are processed in associative memory (Marcus et al. 1992, 1995; Pinker 1991; Pinker and Prince 1988, 1991, 1994; Prasada and Pinker 1993; Jaeger et al. 1996, Ullman 1999). The opposite single-system approach, in its two variations, the connectionist (MacWhinney and Leinbach 1991; Plunkett and Marchman 1991, 1993; Rumelhart and McClelland 1986) and network (Bybee 1985, 1995; Langacker 1987, 1988) approaches, holds that both regular and irregular verbs are processed by one single mechanism in associative memory.

The proponents of the dual- and single-system approaches make opposite predictions about the role of input frequencies in processing English past-tense regular and irregular verbs. According to the dual-system approach, only irregular verbs, which are retrieved from associative memory, will be frequency-sensitive. The single-system approach predicts that frequency will influence the processing of both regular and irregular verbs.

Experimental data on frequency effects in English past-tense inflection are controversial. While some studies demonstrate frequency effects only in irregular verbs (Prasada, Pinker, and Snyder 1990; Ullman 1999) and therefore do not support the single-system view, other studies demonstrate frequency effects for regular verbs as well (Stemberger and MacWhinney 1988; Marchman 1997). One study, which showed frequency effects in regular inflection, measured reaction times in a lexical decision task involving English verbs, nouns, and adjectives, as well as nonce forms. It detected whole-word frequency effects for regularly inflected verbs above the threshold of about 6 per million when stem-cluster frequencies were held constant (Alegre and Gordon 1999).



Thus, the role of frequency effects in English past-tense inflection remains an unresolved issue<sup>4</sup>. However, even if this were not the case, English past-tense inflection with only one regular verb class and the virtually non-existent conjugational paradigm, obviously cannot be readily generalized to other languages with developed inflectional morphology. And indeed, the emerging data on languages with rich verbal morphology indicate that these languages do not exhibit a sharp distinction between regular and irregular verb processing (Ragnasdóttir, Simonsen, and Plunkett 1997; Matcovich 1998; Orsolini and Marslen-Wilson 1997; Orsolini, Fanari, and Bowles 1998; Simonsen 2000). Both developmental and adult data on past tense processing in Italian demonstrate the effects of phonological similarity even in the Conjugation I class, considered to be a regular and default class (Matcovich 1998). Two developmental studies of child first language (L1) acquisition of verbal morphology, one in Norwegian and Icelandic and the other in Italian – languages with complex verbal morphology – recorded the influence of both type and token frequencies on their subjects' responses (Ragnasdóttir, Simonsen, and Plunkett 1997; Matcovich 1998). For Norwegian and Icelandic, this influence was manifested in generating past participles of both strong (irregular) and weak (regular) verbs; in Italian, it was manifested across the verb conjugation classes. The results of these studies, which assessed the influence of input frequencies through the rates of overgeneralization, are in conflict with the predictions made by the proponents of the dual-system approach. Consequently, the application of the modular approach, which assumes a sharp distinction between regular and irregular processing, needs to be reconceptualized in regard to languages with complex morphology.

Given the fact that languages with complex morphology do not easily lend themselves to the same type of analysis as English, one should probably seek a more flexible theoretical position in dealing with them. And indeed, in addition to two strong and consistent positions held by the proponents of the dual- and single-system approaches, there exist several attempts to create models of morphological processing which would integrate both of these approaches. The dual-route model (Baayen et al. 1997) claims that symbolic rule application and search in associative memory proceed simultaneously and whichever route is faster wins the "competition." Another model, which posits the parallel activation of the two systems, the memory system and the rule system, in morphological processing, is the mental model of morphology (Ullman 2000). This model views the memory and rule systems as domain-general, and therefore departs from the classical modular approach based on the idea of domain-specificity. However, it postulates modularity within morphology itself.

An alternative solution to the sharp regular-irregular distinction, which may be more appropriate when dealing with languages with complex morphology, would be to range the verb types on a continuum of morphological complexity and then look for possible dissociations in verbal processing. In fact, research on languages with rich verbal morphology has made attempts to account for the role of morphological complexity in verbal processing. For example, the study of L1 acquisition of Norwegian and Icelandic makes a claim that morphological complexity influences developmental rates. Icelandic has more complex verbal morphology than Norwegian. In accordance with this fact, the Icelandic children were delayed in relation to the Norwegian children at age 4 on the strong verbs (Ragnasdóttir, Simonsen, and Plunkett 1997). A series of studies on morphological processing in German, a language with more complex inflectional morphology than English, raises another important issue. These studies demonstrate that the most likely, if not the only candidate for symbolic rule application is the default class, while the status of other regular non-default classes in morphological processing remains highly controversial (Clahsen 1999)<sup>5</sup>. The polemics reviewed above focus entirely on the formal aspects of the processing of verbal morphology and do not address the role of inherent verb semantics in contextualized morphological marker application. Research on the acquisition of temporal-aspectual and form-meaning relations emphasizes the importance of inherent verb semantics in the use or nonuse of morphological markers in speaking. Moreover, recent studies report on interaction between formal and semantic properties in morphological rule learning. Thus, it appears that primitive conceptual-semantic notions such as stativity, durativity, and telicity mainly affect the processing of regular rule-based morphology, and not irregular morphology in L2 acquisition of English (Housen 2002: 107; Housen 2003: 188).

### **3. Goals of the study and the verbal system of Russian**

This study investigates the processing of verbal morphology in Russian, a language with numerous verb classes and complex conjugational paradigms, by American formal learners of Russian<sup>6</sup>. It compares the results obtained for second language (L2) learners with the baseline data from adult Russian native speakers, and addresses the following questions:

1. Does explicit instruction in complex morphological rules<sup>7</sup> result in the successful learning of these rules as reflected on pencil-and-paper tests?

2. Does explicit instruction on verb conjugation facilitate the development of native-like verbal processing strategies in L2 learners?
3. What is the role of input frequencies in L2 processing of complex verbal morphology?

The study introduces the parameter of the complexity of paradigm, and *a priori* establishes a hierarchy of complexity for the verb classes it includes. It uses type frequencies in standard Russian and input frequencies specifically computed for the L2 learners taking part in the experiment. It is important to note that these learners were shaped by a communicative classroom with focus on form instruction. The course is video-based and includes various communicative activities in addition to explicit formal training in verb conjugation.

What will follow is a short introduction to Russian verb conjugation based on the one-stem verb system developed by Jakobson and his followers (Jakobson 1948; Townsend 1975; Davidson, Gor, and Lekic 1996). This is not the only system describing Russian verb conjugation, and there is an alternative description endorsed by the Russian Academy of Sciences. There is no research to date demonstrating the psycholinguistic validity of either system, and we have chosen the one-stem system for two main reasons:

1. This description allows the generation of all forms of all Russian verbs, with the exception of a dozen truly irregular ones by the application of a set of rules. These sets of rules are different for the different verb classes (and subclasses).
2. The one-stem system was used in the instructional setting for the group of American learners taking part in the matching experiment.

According to the one-stem description, Russian has 11 verb classes, each with its own suffix (also called morphological marker or verb classifier). The eleventh class has a zero suffix, and is subdivided into smaller subclasses depending on the quality of the root-final consonant and the root vowel. This is a small class, especially given the variety of conjugational patterns it includes, and there are less than 100 basic stems in it<sup>8</sup> (Townsend 1975). The remaining 10 suffixed classes are identified by the suffix. These are the following classes: -aj-, -ej-, -a-, -e-, -i-, -o-, -ova-, -avaj-, -nu- (including the “disappearing -nu-”), and -zha. The suffix determines all the parameters of the conjugational paradigm, which include:

1. Conjugational type (type of endings), 1<sup>st</sup> or 2<sup>nd</sup>;
2. Consonant mutations (mutations of the root-final consonant);
3. Stress shift (specific patterns in the -a-, -e-, and -i- classes);
4. Suffix alternations (-ova- alternates with -uj-, -avaj- alternates with -aj-, and -(nu)- disappears under specific conditions).

In addition, there are vowel alternations in zero-suffixed stems: “o” in the (o)j-stem alternates with “y.”

Before taking a closer look at the stems<sup>9</sup> chosen for our experiment, we will review some facts about the morphological processes taking place in verbal conjugation. When the endings are added to the stem (which includes the optional prefix, the root, and the suffix), an automatic truncation rule works at the juncture of the stem and the ending. If the stem ends in a vowel and the ending begins in a vowel, the first vowel is truncated. The same is true for the consonants: the first one is deleted. Past tense endings begin with a consonant, and non-past tense endings begin with a vowel; therefore, stem-final vowels will be deleted in the non-past tense forms, and consonants will be deleted in the past tense forms.

The morphological processes in two Russian verbs, *chitat'* ‘to read’, and *pisat'* ‘to write’, will illustrate this description. Despite the fact that their infinitives look similar, they belong to different stems and have different conjugational patterns. The verb *chitat'* belongs to the -aj- class, and its stem *chit-aj-* ends in a consonant. In the past tense (and the infinitive as well), the -j- is truncated before consonantal endings:

*chit-aj-* + *-l* = *chital* ‘he read’

In the non-past tense, vocalic endings are simply added to the stem:

*chit-aj-* + *u* = *chitaju* ‘I read’

The verb *pisat'* belongs to the -a- class, with its stem ending in a vowel, and the past-tense consonantal ending is added to the stem:

*Pis-a-* + *-l* = *pisal* ‘he wrote’

In the non-past tense, vowel truncation takes place:

*Pis-a-* + *-u* = *pishu* ‘I write’

Note the consonant mutation “s”–“sh”, which in the -a- stem occurs throughout the non-past paradigm.

Table 1. Morphological processes in the stems included in the experiments

Verb classes	-aj-	-a-	-ej-	-e-	(i)j-	-i-	-ova-	-avaj-	(o)j-
Consonant truncation before conson. endings	✓		✓		✓			✓	✓
Vowel truncation before vowel endings		✓		✓		✓	✓		
Consonant mutation		✓		✓		✓			
Stress shift <sup>10</sup>		✓		✓		✓			
Suffix alternation							✓	✓	
Vowel alternation									✓

Table 1 lists the morphological processes or “rules” shaping the conjugational patterns of the stems chosen for the experiment and includes automatic consonant or vowel truncation, which occurs at the juncture of the stem and the ending<sup>11</sup>. Note that the table does not list the stems, which are not part of the experimental material. The reasons for the actual choice of the stems are provided in section 4 below. The most straightforward way to assess the complexity of paradigm for each verb class is to add up all the morphological processes occurring in this class. Thus, the -aj-, -ej-, and (i)j- stems have only one rule, that of consonant deletion, in their paradigm. The -a-, -e-, and -i- stems have three. The -ova-, -avaj-, and (o)j- stems have two. However, this mechanical computation does not take into account the relative complexity of the individual rules. Consonant and vowel truncation are automatic processes, which take place in every conjugational pattern. Consonant mutation and stress shift occur only in the -a-, -e-, and -i- stems; therefore, they are less common, and involve more complex rules. Suffix alternation in the -ova- and -avaj- stems is even more marginal in the Russian verbal system. The vowel alternation occurring in the 5 (o)j- stems places this class in the exceptions category.

#### 4. Experimental material

Type and token frequencies (whole-word and stem-cluster) were shown to influence verbal processing in both adult and child native speakers. But while

adult native speakers potentially have full access to type and token frequencies, formal L2 learners with lower proficiency in L2 have limited access to input frequencies in the target language. A beginning classroom typically exposes learners to most verb classes (types), but the relative size of these classes (type frequency) is not available to the learners, and the frequency of use of individual verb classes may differ substantially from the one found in native Russian. Likewise, token frequencies of individual verbs used in a highly structured situation of learning and a controlled classroom setting do not reflect the ones found in native speech. As a result, L2 learners may develop an interlanguage (IL) system based on verb classes of a more uniform size than the classes in the native language and with non-native token frequencies of individual verbs. Therefore, one can hypothesize that native input frequencies will affect non-native verbal processing indirectly, only to the extent that they are reflected in the actual L2 input frequencies.

*Table 2.* Type frequencies of the verb classes included in the experiments:  
Native and second language input

Verb classes	-aj- Productive	-a- Productive	-ej- Productive	-e- Productive	(ij)- Productive	-i- Productive	-ova- Productive	-avaj- Productive	(oj)- Productive
Russian language Type frequency	11814	940	608	328	160	7019	2816	94	98
Russian language Number of unprefixated stems		Appr. 60 stems		Appr. 50 stems	7 stems			3 stems	5 stems
Input to L2 learners Type frequency	55 (86) <sup>12</sup>	14 (24)	0 (4)	8 (12)	3 (3)	52 (80)	13 (34)	2 (7)	2 (5)
Input to L2 learners Number of uses	4333	1298	12	782	239	4546	555	273	158

Accordingly, the study uses its own frequency counts, which were done with the assumption that the frequencies found in the instructional materials used in first-year Russian would be the best approximation available of the

input frequencies to which our subjects were exposed<sup>13</sup>. The type frequencies and the number of uses of all the verbs were computed for two volumes of the textbook and two volumes of the workbook, which are part of the instructional package *Live from Moscow!* (Davidson, Gor, and Lekic 1996) that was used in first-year Russian. The counts included not only all of the verbs present in the books, but also the verbs in exercises that the students had to generate themselves. For example, if the assignment was to say where the student eats his/her breakfast, lunch, and dinner, the verb “to eat” was counted 3 times in the 1<sup>st</sup> person singular non-past tense. The type frequencies found in the input to the learners were compared with the data on the Russian language (Townsend 1975; Zalizniak 1980).

Table 2 contains information about the type frequencies of the stems included in the experiment and about their productivity. In the first row corresponding to the Russian language, the numbers in each column represent the results of the verb counts, which we performed on the *Grammatical Dictionary of the Russian Language* (Zalizniak 1980) with approximately 100,000 entries. These counts contain all the verbs belonging to a particular conjugation class and include the prefixed and reflexive verbs. The second row provides the number of unprefixed stems for the small unproductive classes (based on Townsend 1975 and Davidson et al. 1996). The next two rows contain two types of data on input frequencies obtained for L2 learners taking part in Experiment 2. The third row shows the L2 type frequencies, and the fourth row the data on the frequency of use, which were computed by adding all the occurrences of all the verbs belonging to each stem<sup>14</sup>. In the counts based on first-year Russian instructional materials, all the prefixed and reflexive verbs were computed as separate items. Therefore, the counts in the first row for Russian language and the third row for L2 speakers are based on the same criteria. According to Joan Bybee (1995), type frequency contributes to the productivity of a given schema<sup>15</sup>. And indeed, the type frequencies and productivity of the stems used in the study confirm the prediction made by Bybee (1995) that the patterns (schemas) with high type frequency are productive. One can expect the conjugational patterns of the productive classes with high type frequency to be generalized more often than the patterns for the low type frequency classes.

The experimental material included 3 pairs of stems, which have a similar past tense (and infinitives as well), but have different conjugational patterns in the non-past tense:

-aj- and -a-  
 -ej- and -e-  
 (i)j- and -i-

The stem is not recoverable in the past tense because the “j” is truncated; therefore the speakers need to “guess” the underlying stem to conjugate the verb in the non-past tense. The experiment aims at establishing which conjugational patterns will be generalized. Of these 6 stems, three belong to the high type frequency productive classes -aj-, -ej-, and -i-. One can expect the conjugational patterns of the productive classes with high type frequency to be generalized more often than the patterns for the low type frequency classes. Therefore, one can predict that the -aj-, -ej-, and -i- patterns will be generalized to the -a-, -e-, and (i)j- classes. At the same time, the three stems ending in “j,” -aj-, -ej-, and (i)j-, have less complex conjugational patterns, and if the complexity of paradigm plays a role in verbal processing, these stems should be more generalizable. One can easily see that there is a conflict between two predictions for the -i- and (i)j- stems. From the point of view of the complexity of paradigm, the (i)j- pattern should be generalized. But the (i)j- pattern occurs only in 7 stems; therefore based on type frequency, the -i- pattern should be generalized.

The next two stems included in the testing material, -ova- and -avaj-, have similar conjugational features – they show suffix alternations in the non-past tense: -ova- alternates with -uj-, and -avaj- alternates with -aj-. For such stems, the past tense form contains sufficient morphological information (morphological cues) for the speakers to be able to identify the stems. However, these classes differ radically in their type frequencies: the -ova- class has high type frequency, whereas the -avaj- class has only three basic stems. The experiment tests whether the subjects actually pay attention to the morphological cues and produce the suffix alternation expected in these stems, and whether type frequency influences their processing.

The last stem, or more exactly, the subclass of zero-suffixed stems, (o)j-, has a very special feature: alternation of the root vowel in the past tense: “o” alternates with “y”. The vowel “y” does not occur in any of the suffixes, and the past tense form of such verbs sounds unusual. This stem was included to test whether the presence of the vowel “y” serves as a cue for the low type frequency (o)j- stems.

The study included 3 experiments; Experiments 1 and 2 involved American learners, and Experiment 3, which matched Experiment 2, native Russian speakers.



## **5. Experiment 1 with American learners (written)**

### 5.1. Experimental procedure

Experiment 1 was a paper-and-pencil test, which tested the knowledge of the conjugational rules, since the conjugational pattern was recoverable from the provided stimuli<sup>16</sup>. This experiment aimed at establishing the baseline for the oral Experiment 2 with unrecoverable stems. The data for Experiment 1 were collected from 15 volunteer students at the University of Maryland, College Park in the middle of their third semester of Russian. The testing material consisted of 46 real Russian verbs belonging to 9 stems (see the Appendix for the list of verbs)<sup>17</sup>. The experiment consisted of 2 parts. In the first part, the stimulus verb was provided as a basic stem<sup>18</sup>, and the subjects were asked to generate 3 non-past tense forms: the 1<sup>st</sup> and 2<sup>nd</sup> person singular, and the 3<sup>rd</sup> person plural. In the second part, 2 forms of the stimulus verb were provided: the past tense plural and the non-past 3<sup>rd</sup> person plural. The subjects had to generate 2 forms: the 1<sup>st</sup> and 2<sup>nd</sup> person of the non-past tense. In both parts of the test, the conjugational pattern was recoverable from the provided stimuli. In addition to establishing the baseline for the oral Experiment 2 with unrecoverable stems, the goal of Experiment 1 was to determine, which condition, “stem” (the stimulus verb provided as a basic stem) or “forms” (the stimulus verb provided as two verb-forms), facilitates verbal processing to a greater extent.

### 5.2. Results of experiment 1

The experiment measured two parameters of the subjects' responses: the rates of correct stem recognition (which ignored errors in consonant mutation and conjugation type), and correct response rates for each verb class. All of the 15 students' responses to the set of 46 verbs were transcribed, and the scores for all individual verbs grouped by the stem were entered in a table as numbers of verbs that were conjugated as belonging to one of the verb classes included in the experiment. The last column, “Other,” was reserved for the responses that did not follow the paradigm for any of the stems used in the initial set of verbs. By averaging these data we computed the percentages of stem recognition as well as the rates of generalization for each verb class, which are represented in Table 3. The experiment tested the students' knowledge of the Russian conjugational system, and more specifically, of its one-stem description. Since most of the verbal

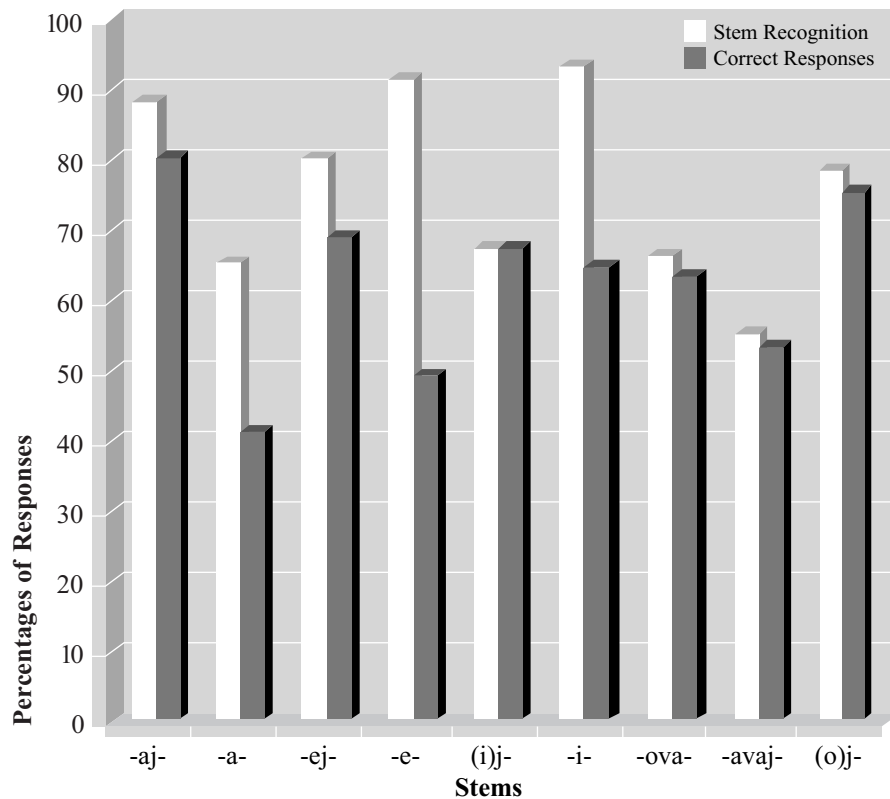


Figure 1. Rates of stem recognition and correct responses in the written experiment (American learners)

stimuli were unfamiliar to the subjects, the written experiment revealed not their knowledge of the conjugational paradigms of individual verbs, but their ability to process the novel verbs belonging to different classes. The results of Experiment 1 demonstrated that the beginning American learners were reasonably good at stem recognition. For most stems the rate was above 60%, and for 5 stems it was above 75%. Figure 1, which combines the results for both parts of the experiment, the “stem” and “forms” conditions, compares the rates of stem recognition and correct responses. The percentage of correct responses was predictably lower. The greatest discrepancies are found in the -a-, -e-, and -i- stems, which have complex conjugational patterns including consonant mutations and stress shifts, and which belong to different conjugation types, the latter fact representing another source of confusion for beginning learners.

Table 3. Distribution of responses in experiment 1 with American learners: Stem recognition (written)

Stimuli	Responses									
	-aj-	-a-	-ej-	-e-	(i)j-	-i-	-ova-	-avaj-	(o)j-	Other
-aj-	<b>88.0</b>	3.6			0.4			1.3		6.7
-a-	28.0	<b>64.9</b>							0.4	6.7
-ej-			<b>80.4</b>	14.3	1.3					4.0
-e-	0.4	0.8	4.7	<b>91.0</b>						3.1
(i)j-		6.7			<b>67.0</b>	13.3				6.7
-i-	0.7	1.3			1.0	<b>93.0</b>		2.0		2.0
-ova-	16.9	3.1					<b>65.8</b>	0.4	3.1	10.7
-avaj-	40.6	0.6	0.6					<b>55.0</b>		3.2
(o)j-						4.0			<b>86.7</b>	5.3

The “stem” condition required the use of the most analytical procedures – the subjects had to apply the conjugational pattern to the abstract representation of the stimulus verb, its basic stem. The “forms” condition provided two real verb-forms, which, depending on the specific stem, could contain the information about the conjugation type, the conjugational pattern (“Vowel+j” in the -aj-, -ej-, and (i)j- stems versus “Vowel” in the -a-, -e-, and -i- stems), consonant mutations, suffix, and vowel alternations. The comparison of the results for these two conditions established that, overall, the “forms” condition produced higher rates of stem recognition than the “stem” condition. The Z-test was performed for 8 stems<sup>19</sup>, and on 5 of them the “form” condition scored significantly higher at the 0.05 confidence level. Only in the -i-, -e-, and (o)j- stems were there no significant differences between the two conditions.

## 6. Experiment 2 with American learners (oral)

### 6.1. Experimental procedure

The same group of 15 volunteer students at the University of Maryland took part in Experiment 2. It was conducted orally and individually with each subject, and recorded on audiotape. The subjects met with the experimenter and received the printed version of the test, which included written instruc-

tions. The experimenter read aloud all the sentences carrying the verb in the past tense and the question designed to trigger the use of non-past forms by the subject. A warm-up containing the verbs not included in the main testing material preceded the main part of the experiment.

The testing material consisted of 48 real Russian verbs, which served as prototypes for the nonce verbs used in Experiment 3 with native speakers of Russian. These 48 verbs belonged to the following verb classes and subclasses of non-suffixed stems (based on Jakobson's one-stem verb system): -aj-, -ej-, (i)j-, (o)j-, -a-, -e-, -i-, -ova-, -avaj- (see the Appendix for the list of verbs). The number of verbs in each class varied from 2 for very small (i)j- and (o)j- subclasses to 6-8 in other classes. Half of the students received the test with a different order of verb presentation – the last 24 verbal stimuli were moved to the beginning of the test in order to control for the fatigue factor. The verbal stimuli were in the past tense plural form. The subjects were asked to generate the non-past 3<sup>rd</sup> person plural and 1<sup>st</sup> person singular forms of the verbal stimuli. All the verbs were embedded in simple carrying sentences, which, together with follow-up questions, formed a quasi-dialogue:

Experimenter: Yesterday they \_\_\_\_\_. And what are they doing today?  
 Subject: Today they \_\_\_\_\_<sup>20</sup>.  
 Experimenter: And you?  
 Subject: Today I \_\_\_\_\_.

This elicitation technique is based on the adaptations of the instrument developed by Bybee and Slobin (1982) used in the studies of child L1 acquisition of Norwegian, Icelandic (Ragnasdóttir, Simonsen, and Plunkett 1997; Simonsen 2000), and Italian (Matcovich 1998).

After the oral part was completed, students received the list of all the verbs included in the main part of the test and were asked to check off all the verbs that they knew. Most of the verbs in the testing material were unfamiliar to the learners, but we decided to include several verbs from the active 1<sup>st</sup> year vocabulary frequently used in the classroom to make the task psycholinguistically more authentic for the beginning learners of Russian. In order to control for the familiarity factor, we performed all the computations separately for 25 verbs that no more than 2 students identified as familiar. Since the results for these 25 unfamiliar verbs showed exactly the same tendencies as the entire sample, we will discuss the data on all the 48 verbal stimuli.

6.2. Results of experiment 2

This experiment also computed the rates of stem recognition and correct responses. Table 4 contains the rates of stem recognition collapsed for each stem. The results obtained for the -aj- and -a- stems indicate that L2 learners identified and generalized the high type frequency -aj- pattern. The next pair of stems, -ej- and -e-, manifests a weaker tendency to identify and generalize the -ej- pattern. In the last, “problematic,” pair of stems, (i)j- and -i-, the tendency to generalize the (i)j- pattern is even weaker than in the two preceding stems. At the same time, the -i- pattern had low generalizability as well. Overall, the non-native processing of the “paired” stems shows the following tendency: the -aj- pattern demonstrates high generalizability, while the -ej- and (i)j- patterns are used more often in the responses to the appropriate stems than to their counterpart stems without

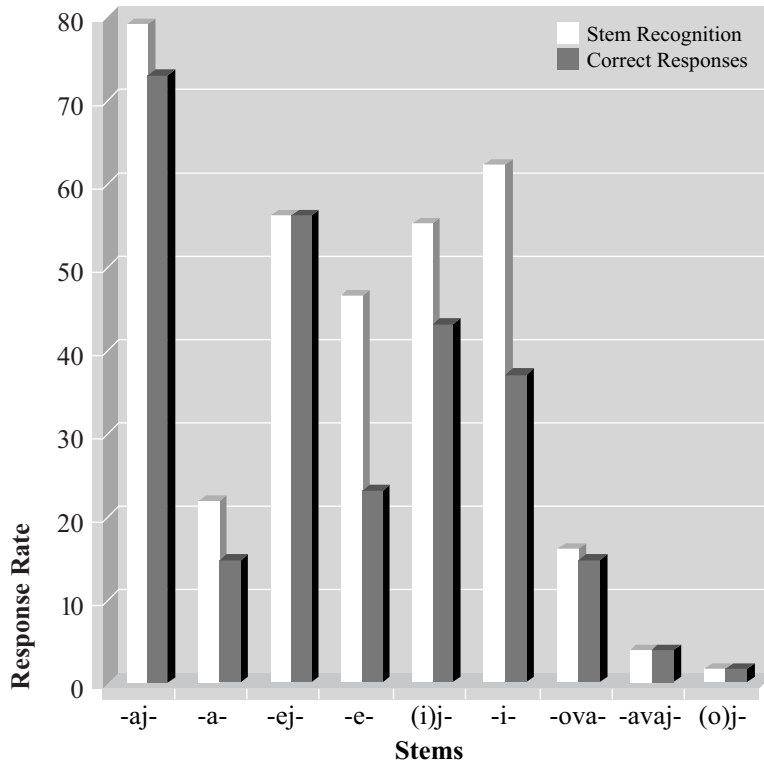


Figure 2. Rates of stem recognition and correct responses in the oral experiment (American learners)

the -j-. Thus, in the last two pairs of stems, the American learners proved to be reasonably efficient at “guessing” the underlying unrecoverable stem.

L2 learners experienced difficulties with the identification of the -ova- and -avaj- suffixes (or possibly, with the application of the pattern with suffix alternation). The low type frequency (o)j- stem was very poorly identified. The responses to this stem showed an interesting feature: 1/3 of the (o)j- verbs were conjugated using the \*(y)j- pattern illegal in Russian. Figure 2 compares the rates of stem recognition and correct responses in this experiment and reveals that the differences are the greatest for the -e- and -i stems with complex conjugational paradigms.

*Table 4.* Distribution of responses in experiment 2 with American learners: Stem recognition (oral)

Stimuli	Responses									
	-aj-	-a-	-ej-	-e-	(i)j-	-i-	-ova-	-avaj-	(o)j-	Other
<b>-aj-</b>	<b>79.0</b>	12.0			2.2			0.8		6.0
<b>-a-</b>	66.0	<b>22.0</b>			0.5	0.5				11.0
<b>-ej-</b>	3.9	18.2	<b>56.0</b>	9.3	1.1	0.5				11.0
<b>-e-</b>	5.2	1.9	30.0	<b>46.6</b>	4.3		1.0			11.0
<b>(i)j-</b>	1.6	15.0		1.6	<b>55.0</b>	22.0			1.6	3.2
<b>-i-</b>	10.0	2.5	7.1		7.9	<b>62.0</b>				10.5
<b>-ova-</b>	68.3					6.7	<b>16.1</b>		0.6	8.3
<b>-avaj-</b>	80.6		0.6			3.3	5.8	<b>3.9</b>		5.8
<b>(o)j-</b>			1.7		13.3	1.7			<b>1.7</b>	81.6

## 7. Experiment 3 with Russian speakers (oral)

### 7.1. Experimental procedure

Experiment 3 was conducted with 27 adult Russian speakers at St. Petersburg State University, and provided a native baseline for Experiment 2 with American learners of Russian. The experimental procedure was the same as in Experiment 2. The testing material consisted of 48 nonce verbs, which we created by manipulating the initial segments in the real Russian verbs used with American learners (see the list of verbs in the Appendix).

The aim of modifications was to preserve as much of the phonological shape of the real verbs as possible, but at the same time, to avoid any close resemblance to the real verbs. In most cases, only the initial consonant was modified.

7.2. Results of experiment 3

Table 5 shows the distribution of the rates of stem recognition for this experiment. In the “paired” stems, Russian speakers consistently identified and generalized the nonce verbs derived from the -aj- and -ej- classes. However, the same tendency cannot be observed in the -i- and (i)j- stems. Here the frequent -i- pattern is not dominant, in fact, neither pattern is very active. Native speakers conjugated approximately 1/2 of the -ova- verbs using the -ova- pattern involving the suffix alternation. But they did even less well on the -avaj- stem. Less than 1% of the (o)j-verbs with “y” in the past tense were conjugated as (o)j- stems. At the same time, approximately 1/2 of the (o)j- verbs were conjugated using the non-existent \*(y)j- pattern.

Table 5. Distribution of responses in experiment 3 with native speakers of Russian: Stem recognition (Oral)

Stimuli	Responses									
	-aj-	-a-	-ej-	-e-	(i)j-	-i-	-ova-	-avaj-	(o)j-	Other
-aj-	<b>89.7</b>	0.6	3.1			3.5				3.1
-a-	80.9	<b>11.7</b>	3.1							4.3
-ej-	0.6		<b>73.8</b>	7.4	0.6					17.6
-e-	1.3		60.4	<b>25.5</b>			0.8			12.0
(i)j-			2.8		<b>24.1</b>	16.7	8.3			48.1
-i-	2.6		12.2		30.0	<b>36.6</b>	0.2			18.4
-ova-	40.4					1.6	<b>47.2</b>			10.8
-avaj-	61.1		3.0			0.4	9.3	<b>21.8</b>		4.4
(o)j-	0.9		0.9			6.5			<b>0.9</b>	90.8

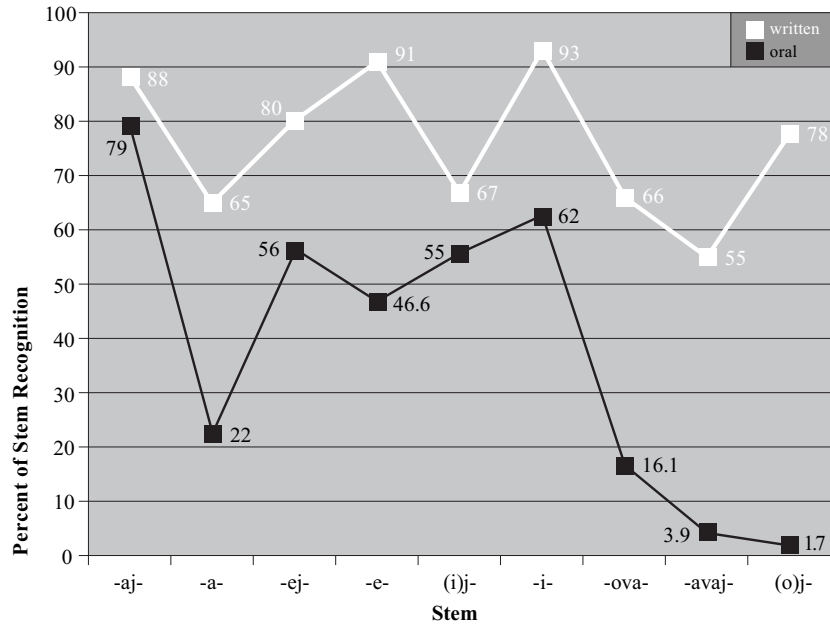


Figure 3. Rates of stem recognition in the oral and written experiments (American learners)

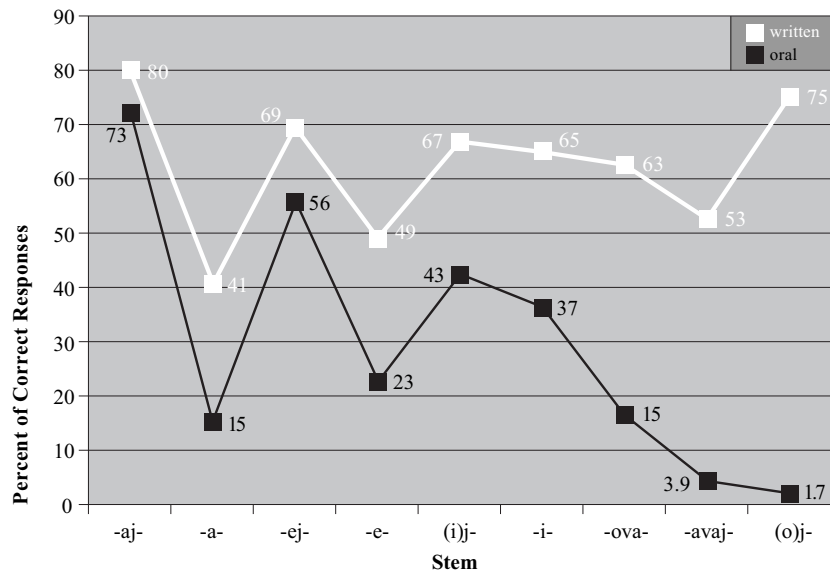


Figure 4. Rates of correct responses in the oral and written experiment (American learners)



## **8. Comparison of the results of experiments 1, 2, and 3 and discussion**

This section will compare and discuss two sets of data, oral and written, for the American learners, and the oral results obtained for the American learners and native speakers of Russian. It will then analyze the experimental evidence on the rates of missed consonant mutations in Experiments 1, 2, and 3.

### 8.1. Comparison of the written and oral results (experiments 1 and 2)

The comparison of the rates of stem recognition and correct responses in Experiments 1 and 2 yielded a fully predictable result: they were higher in the written experiment than in the oral one (see Figures 3 and 4). The Z-test performed on the results established that this difference was statistically significant for all the stems at the 0.05 confidence level. Apparently, two factors contributed to this effect. The most important factor is that in the written experiment, all the stems were recoverable from the provided stimuli, whereas in the oral experiment 6 “paired” stems were unrecoverable. Accordingly, the subjects were performing different tasks – applying their knowledge of the conjugational system in the written experiment and “guessing” based on their knowledge of the conjugational system, which included statistical probabilities, in the oral experiment. The analysis of the data for individual stems reveals that the greatest difference between the written and oral results is observed in the stems containing the morphological cues, -ova-, -avaj-, and especially (o)j-. This is due to the presence of the target conjugational pattern in the provided stimuli for these stems, in the “forms” condition for all three of them, and in both conditions for the (o)j- stem.

### 8.2. Comparison of non-native and native processing (experiments 2 and 3)

Figure 5 compares the rates of stem recognition in the two oral experiments with American learners and native speakers of Russian. The discussion will focus, first, on the similarities, and then on the differences in the obtained results.

Both groups of subjects identified and generalized the productive -aj- and -ej- patterns. Both the American and Russian speakers created the \*(y)j- pattern illegal in Russian in response to the (o)j- stems. Thus, both non-native and native speakers relied heavily on the “Vowel + j” pattern, apparently

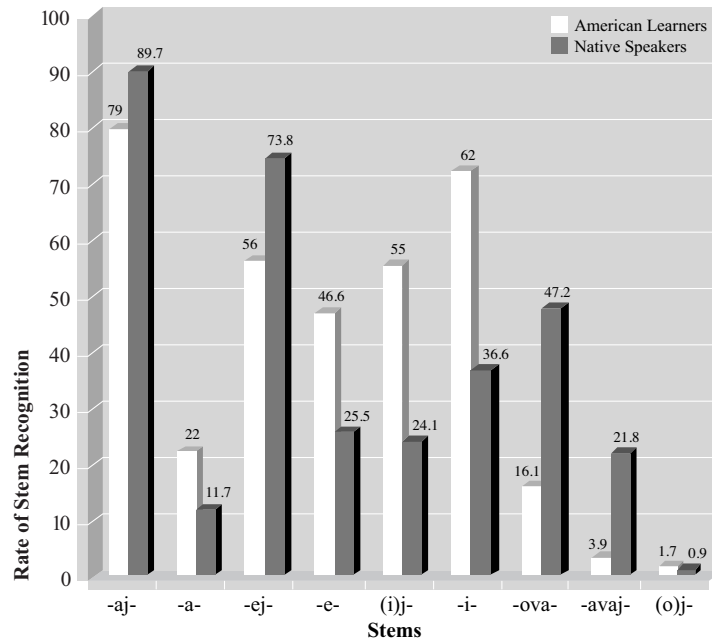


Figure 5. Rates of stem recognition in American learners and Russian native speakers

the default pattern in Russian. The results obtained for the -aj-/-a- and -ej-/-e- pairs fully confirm the prediction that the conjugational pattern of high type frequency verbs will be generalized to low type frequency stems. At the same time, these results are not in conflict with the predictions based on the complexity of paradigm. We will return to the L2 data regarding the -ej-/-e- stems later in this section. However, the results for the -i- and (i)j- stems do not show the same tendency. Here the frequent -i- pattern was competing with a less complex (i)j- pattern. Since the -i- class is productive and has high type frequency, while the (i)j- subclass of zero-suffixed verbs includes only 7 basic stems, it was expected that the -i- pattern would be generalized to the (i)j- stems. However, the rate of generalizations of the -i- pattern was relatively low. There are two possible interpretations for this effect:

1. In individual stems the complexity of paradigm factor overrides the frequency factor.
2. The overall pattern of responses in the experiment suggests that both groups of subjects favour the default rule “recover the j,” regardless of

the stem. Consequently, the low generalization rates for the -i- stem in comparison with the -aj- and -ej- stems could reflect the differences between default and non-default processing rather than regular and irregular processing.

The -ova- class has high type frequency and is productive, while the -avaj- class includes only 3 stems. Therefore, the fact that the -ova- marker worked better as a cue to the conjugational pattern than the -avaj- marker for both groups of subjects, confirms the role of frequency in processing novel verbs. However, the -ova- cue has limited efficiency and does not automatically trigger the suffix alternation. As for the (o)j- class, which includes only 5 stems, the results indicate that neither group of speakers established analogies with this class in processing the verbal stimuli.

At the same time, the Z-test showed significant differences in the rates of stem recognition in L2 and L1 processing for all the stems, except the (o)j- stem at the 0.05 confidence level. American learners had significantly lower rates of stem recognition for the -aj-, -ej-, -ova-, and -avaj- stems, as well as significantly higher rates for the -a-, -e-, -i-, and (i)j- stems (see Figure 5). These differences suggest that the beginning American learners relied on the default “Vowel+j” pattern less often, and generalized less “regular” patterns more often than the native Russians.

One of the goals of this study was to compare the input frequencies to L2 learners and native speakers of Russian, and determine if the differences between them can account for the differences in the results obtained for the two groups of subjects. While the type frequencies of the different verb classes obtained for non-native input generally reflect the ranking of classes in Russian language, there are significant differences (see Table 2):

1. Quantitative differences between the verb classes are less salient in the L2 input.
2. The -ej- stem is practically not represented in the L2 input (it occurs only in the passive vocabulary).

Given the leveling of differences in class size in the non-native input, one could expect fewer generalizations of high type frequency classes to small classes and less reliance on the default patterns in L2 processing. And indeed, L2 speakers tended to generalize the -aj- pattern less than the native speakers. The most striking result, however, is that L2 speakers consistently identified and generalized the -ej- pattern, which was very poorly represented in the L2 input, and the (i)j- pattern, which has low type frequencies

both in native Russian and the input to L2 learners. Here, L2 learners were relying not on input frequencies, but rather on the default “Vowel+j” pattern in Russian. At the same time, they were avoiding the more complex conjugational pattern. Thus, the American learners were efficient in their use of the native processing strategy – generalization of the default pattern. At the same time, they used the default pattern less often than the native speakers, probably at least in part due to the leveling effect in the input they had received.

### 8.3. Consonant mutations (experiments 1, 2, and 3)

The results of both oral experiments suggest that neither non-native nor native processing necessarily trigger all the rules shaping the conjugational pattern for a particular paradigm. It appears that speakers can single out and apply individual discrete rules, but the whole pattern is not always activated. Thus, one such rule (“recover the j”) was generalized in cases of ambiguity, and its application even resulted in the creation of the \*(y)j-pattern illegal in Russian. One way to test this claim on additional data was to look at the pattern of consonant mutations. The rates of stem recognition did not take into account the errors in consonant mutation, and we analyzed the rates of missed mutations for the -a-, -e-, and -i- stems separately for the oral non-native, written non-native, and oral native data.

*Table 6.* Rates of Missed Consonant Mutations in Experiments 1, 2, and 3

Stems	-a-	-e-	-i-	Average
American Oral	41.4	58.3	34.6	43.9
American Written	19.6	36.7	37.3	29.3
Russian Oral	65.0	63.0	47.4	56.1

Table 6 demonstrates that the rates of missed mutations were high, with native Russians scoring the highest, and American learners scoring the lowest on the writing test. According to the Z-test, the differences between all the three average figures were statistically significant at the 0.05 confidence level, however, the difference between the American and Russian oral data is borderline significant. It still remains to be explained why the Russian speakers omitted slightly more mutations than the American learners.

The first possible explanation involves the use of the avoidance strategy documented in SLA research: because L2 learners are not confident with the application of the mutation rules, they use the forms requiring mutation less frequently. But this explanation is not borne out by the data, since the American learners used the -a-, -e-, and -i- conjugational patterns more often than the Russian speakers. Based on these results, one may speculate that the paradigm is more dissociated in L1 processing, and that L2 learners are more consistent in applying the entire conjugational pattern.

## **9. Conclusions**

This study analyzed the role of input frequency (type frequency), the complexity of paradigm, and morphological cues in L2 and L1 processing of Russian verbal morphology. It demonstrated that American learners approached the task of generating Russian verbs in the same way as Russian native speakers. Both groups of speakers dealt with morphological complexity and relied on morphological cues in similar ways. Both L2 learners and L1 speakers identified and generalized the default “Vowel+j” pattern. The high rates of missed consonant mutations and the application of the rule “recover the j” to an inappropriate verb class indicate that the rules constituting the conjugational paradigm are not necessarily applied in a set. The type frequencies of the verb classes influenced both non-native and native verbal processing – high frequency conjugational patterns were more readily generalized to other classes. Also, the morphological cues worked better in the processing of high frequency classes. However, in the task, which required the generation of novel verb forms, the complexity of paradigm overrode the frequency factor. At the same time, there are significant differences in the rates of L2 and L1 use of the individual conjugational patterns – L2 learners relied on the default pattern less and generalized the less “regular” patterns more than L1 speakers, which is to a large extent due to the differences in the L2 and L1 input frequencies. In the small subset of the Russian verbs the beginning learners of Russian were exposed to, the differences between the type frequencies of the verb classes were much less salient than in the Russian language as spoken by native speakers.

The obtained results shed new light on the discussion between the proponents of the dual- and single-system approaches to morphological processing. On the one hand, the role of type frequencies both in L2 and L1 processing is an indication that symbolic rule application, which is completely

independent of input frequencies, can hardly be found in the processing of Russian complex morphology. On the other hand, the generalization of the “Vowel+j” pattern to inappropriate verb classes is an indication that the processing of Russian verbal morphology is not entirely dependent on phonological similarity or morphological cues. Thus, neither the dual- nor single-system approach account for all the aspects of the reported data. A hybrid model combining these two approaches seems to be in order. Such a model integrating both approaches, which has a better fit to the reported data, is actually being developed (Gor 2004).

Therefore, the study has demonstrated that in processing complex verbal morphology, L2 speakers could successfully apply the pedagogical rules taught explicitly in the focus on form classroom. In addition to that, structured exposure to the target verbal system led to the development of native-like processing strategies in L2 learners – similarly to native speakers, they generalized the default pattern to other verb classes. L2 speakers relied on their knowledge of the statistical probabilities (type frequencies of the verb classes), which was shaped by the input they had received in the classroom. The role of input frequencies demonstrated in this study emphasizes the importance of the statistical characteristics of the language used in the classroom for the internalization of the target language system and for the development of native-like processing strategies in L2 learners. And finally, the results of this study highlight the positive outcome of focus on form instruction, which combines a meaning-based approach with explicit explanations and practice in the formal aspects of linguistic processing.

### **Acknowledgements**

The work was partly supported by The Russian Scientific Foundation for Research in Humanities (grant 04-04-00083) and by The Russian Foundation for Basic Research (03-06-80068).

## Appendix

### The Russian Verbs and Matching Nonce Verbs Used in the Experiments

(In some cases both simple and progressive forms are possible in translating Russian forms.)

Real Russian verbs	Nonce Verbs	Real Russian verbs	Nonce Verbs
<b>-aj-</b>	<b>-aj-</b>	<b>(i)j-</b>	<b>(i)j-</b>
plAvali (they) swam	klAvali	pIli (they) drank	kIli
pr'Adali (they) moved their ears	tr'Adali	gnIli (they) rotted	bnIli
kArkali (they) croaked	pArkali		
igrAli (they) played	kidrAli	<b>-i-</b>	<b>-i-</b>
gul'Ali (they) walked	tul'Ali	l'ubIli (they) loved	d'ubIli
lAzali (they) climbed	rAzali	proslIli (they) asked	trosIli
		vozIli (they) carried	gozIli
<b>-a-</b>	<b>-a-</b>	plAvili (they) swam	klAvili
pisAli (they) wrote	kisAli	krivIli (they) twisted	tlivIli
mAzali (they) spread	vAzali	glAdili (they) stroked	grAdili
plAkali (they) cried	trAkali	razIli (they) hit	gazIli
klIkali (they) called	tlIkali	mutIli (they) stirred up	lutIli
sYpali (they) strewed	tYpali		
pr'Atali (they) hid	ml'Atali	<b>-ova-</b>	<b>-ova-</b>
		risovAli (they) drew	lisovAli
<b>-ej-</b>	<b>-ej-</b>	diktovAli (they) dictated	biktovAli
umEli (they) were able	ugEli	ekzamenovAli (they) examined	vremenovAli
zhalEli (they) felt sorry	talEli	gazovAli (they) speeded up	bazovAli
krasnEli (they) turned red	plasnEli	prObovali (they) tried	slObovali
imEli (they) had	irEli	kantovAli (they) edged	tantovAli
grEli (they) warmed up	drEli		
tlEli (they) smouldered	glEli	<b>-avaj-</b>	<b>-avaj-</b>
		vtstavAli (they) stood up	vtlavAli
<b>-e-</b>	<b>-e-</b>	prodavAli (they) sold	udavAli
visEli (they) hanged	bisEli	ustavAli (they) got tired	ispavAli
sidEli (they) sat	fidEli	predavAli (they) betrayed	kledavAli
vIdeli (they) saw	mIdeli	uznavAli (they) recognized	oznavAli
kipEli (they) boiled	tipEli		
zavIseli (they) depended	davIseli	<b>(o)j-</b>	<b>(o)j-</b>
xrapEli (they) snored	shkapEli	mYli (they) washed	zYli
gl'adEli (they) looked	br'adEli	krYli (they) covered	brYli

## Notes

1. Schmidt contends that subliminal learning is impossible: “One convincing demonstration of learning without attention would be enough, but so far there have not been any. Several studies purporting to demonstrate learning without attention are shown to really have demonstrated only a low level of learning associated with a low level of attention” (Schmidt 1995b: X).
2. A special volume, “Focus on Form in Classroom SLA”, introduces the acronym for the term, **focus on form instruction** – FonF instruction (Doughty and Williams 1998).
3. Another issue addressed by empirical research is the positive role of negative feedback including explicit corrections on SLA (see Spada and Lightbown 1993; and Trahey and White 1993).
4. Most studies mentioned above deal with token frequency, or the frequency of individual verbs. For them Bybee, the proponent of the network approach, predicts that high-frequency verbs, both irregular and regular, will have high lexical strength and weak lexical associations with other verbs. As a result, these verbs will be easy to retrieve, but will influence other verbs belonging to the same class to a lesser extent (Bybee 1995: 450). Earlier studies did not take into account the distinction between whole-word and stem-cluster frequencies, which could have caused some of the discrepancies in the obtained results. This study is concerned with type frequencies, or in other words, with the sizes of verb classes. According to Bybee, the type frequency of a given class correlates with its productivity (Bybee 1995: 452).
5. For a detailed discussion of this issue see responses to the article by Harald Clahsen (*Behavioral and Brain Sciences* 22, 1999) and Gor 2004.
6. For an earlier discussion of the oral results on adult L1 and L2 verbal processing presented in this paper, see Chernigovskaya and Gor 2000, and Gor and Chernigovskaya 2001.
7. The term “rule” refers to pedagogical rules, or in other words, explicit explanations provided to the learners in the language classroom. It is not intended to mean symbolic rules used in linguistic descriptions or psycholinguistic rules involved in mental processing.
8. Note that two such subclasses, the (o)j- and (i)j-, are part of the testing material. The subclasses of non-suffixed stems are listed by the stem vowel (in parentheses) and the stem-final consonant.
9. Since the verb class in the one-stem system is defined by the stem, this paper will use the terms verb class and stem interchangeably to refer to individual conjugational patterns. Each of these conjugational patterns is characterized by a type frequency – the size of the class, or the number of verbs that are conjugated using this particular pattern. Concurrently, the term stem will also refer to a specific part of the verb.
10. This paper does not discuss stress shifts in the obtained data.



11. All but two of the verb classes included in the testing material belong to the suffixed stems (classes). Two stems, (i)j- and (o)j-, belong to non-suffixed stems ending in -j and differing by the root vowel (these are the subclasses of non-suffixed stems). For these subclasses, the conjugational pattern is defined by the stem vowel and the root vowel and the root-final consonant.
12. The first figure corresponds to the number of verbs in the active vocabulary, and the second figure (in parentheses) to all the verbs from the active and passive vocabulary combined.
13. The advantage of using the subjects who have completed only one year of instruction with a highly structured set of materials is that the experimenters can be confident that the frequencies computed based on the textbook and workbook truly reflect the input the learners have received. This approach becomes much more problematic with more advanced learners.
14. In discussing the data on the acquisition of French verbs (Guillaume 1927/1973), Bybee concludes that the number of uses influences the rate of generalizations less than type frequency (Bybee 1995: 433). We do not support Bybee's interpretation of the French verb data, which treats all the irregular 3<sup>rd</sup> conjugation verbs together, and provide the frequency of use data for our L2 learners. One can observe that except for the -ova- stem, the ranking of the stems based on type frequency and frequency of use is the same.
15. At the same time, Bybee claims that productivity, regularity, and default are different categories, which overlap in verbal processing only to a certain extent (Bybee 1995).
16. This test was administered 1–2 weeks after the oral one depending on the individual schedule of the subjects. By conducting the written test after the oral one, we wanted to prevent any influence of the training potentially occurring in the course of the written test from impacting the results of the oral test.
17. The stimulus set was the same as in Experiment 2, with the exception of 2 verbs.
18. This abstract notation containing the information about the verbal suffix was used as a pedagogical tool in the classroom.
19. The (i)j- stem was represented in this experiment by only one verb.
20. In Russian, present-tense verb forms are synthetic.

## References

- Alanen, R.  
1995 Input enhancement and rule presentation in second language acquisition. In *Attention and Awareness in Foreign Language Learning*, R. Schmidt (ed.), 259–302. University of Hawai'i at Manoa: Second Language Teaching and Curriculum Center.
- Alegre, M. and P. Gordon  
1999 Frequency effects and the representational status of regular inflections. *Journal of Memory and Language* 40: 41–61.

- Baayen, H., T. Dijkstra and R. Schreuder  
 1997 Singulars and plurals in Dutch: Evidence for a dual parallel route model. *Journal of Memory and Language* 37: 94–119.
- Bybee, J. L.  
 1985 *Morphology: A Study of the Relation Between Meaning and Form*. Philadelphia: John Benjamins Publishers.  
 1995 Regular morphology and the lexicon. *Language and Cognitive Processes* 10: 425–455.
- Chernigovskaya, T. and K. Gor  
 2000 The complexity of paradigm and input frequencies in native and second language verbal processing: Evidence from Russian. *Language and Language Behavior* 3 (2): 20–37.
- Clahsen, H.  
 1999 Lexical entries and rules of language: A multidisciplinary study of German inflection. *Behavioral and Brain Sciences* 22: 991–1060.
- Davidson, D. E., K. S. Gor and M. D. Lekić  
 1996 *Russian: Stage One: Live from Moscow!* Kendall/Hunt Publishing Company: Dubuque, Iowa.
- Doughty, C. and L. Williams  
 1998 Pedagogical choices in focus on form. In *Focus on Form in Classroom Second Language Acquisition*, C. Doughty and J. Williams (eds.), 197–261. Cambridge, UK: Cambridge University Press.
- Gor, K.  
 2004 The Rules and Probabilities Model of Native and Second Language Morphological Processing. In: *Teoreticheskiye problemy jazykoznanija*. Sbornik statej k 140-letiju kafedry obshchego jazykoznanija filologicheskogo fakul'teta Sankt-Peterburgskogo gosudarstvennogo universiteta, [In: *Theoretical Problems of Linguistics*. Papers dedicated to 140 years of Department of General Linguistics of St. Petersburg State University.], L. Verbitskaya and T. Chernigovskaya (eds.), 51–75. St. Petersburg: Philological Faculty of St. Petersburg State University Publishing House.
- Gor, K. and T. Chernigovskaya  
 2001 Rules in processing of Russian verbal morphology. In *Current Issues in Formal Slavic Linguistics*, G. Zybatow, U. Junghanns, G. Mehlhorn and L. Szucsich, (eds.), 528–535. Frankfurt/Main: Peter Lang.
- Ellis, N.  
 1993 Rules and instances in foreign language learning: Interactions of explicit and implicit knowledge. *European Journal of Cognitive Psychology* 5: 289–318.
- Harley, B.  
 1993 Instructional strategies and SLA in early French immersion. *Studies in Second Language Acquisition* 15: 245–259.

- 1998 The role of focus-on-form tasks in promoting child L2 acquisition. In *Focus on Form in Classroom Second Language Acquisition*, C. Doughty and J. Williams (eds.), 156–174. Cambridge, UK: Cambridge University Press.
- Housen, A.  
2002 The development of Tense-Aspect in English as a second language and the variable influence of inherent aspect. In *The L2 Acquisition of Tense-Aspect Morphology* R. Salaberry and Y. Shirai (eds.), 155–197. Amsterdam: John Benjamins.  
2003 A corpus-based study of the L2-acquisition of the English verb system. In *Computer Learner Corpora, Second Language Acquisition and Foreign Language Learning*, S. Granger, J. Hung and S. Petch-Tyson (eds.), 77–116. Amsterdam: John Benjamins.
- Jaeger, J. J., A. H. Lockwood, D. L. Kemmerer, R. D. van Valin Jr., B. W. Murphy and H. G. Khalak  
1996 A positron emission tomographic study of regular and irregular verb morphology in English. *Language* 72: 451–497.
- Jakobson, R. O.  
1948 Russian conjugation. *Word* 4: 155–167.
- Jourdenais, R., M. Ota, S. Stauffer, B. Boyson and C. Doughty  
1995 Does textual enhancement promote noticing? A think-aloud protocol analysis. In *Attention and Awareness in Foreign Language Learning*, R. Schmidt (ed.), 183–216. University of Hawai'i at Manoa: Second Language Teaching and Curriculum Center.
- Krashen, S.  
1977 Some issues relating to the monitor model. In *On TESOL '77*, H. Brown, C. Yorio and R. Crymes (eds.), 144–158. TESOL, Washington, D.C.
- Langacker, R.  
1987 *Foundations of Cognitive Grammar, vol. 1. Theoretical Prerequisites*. Stanford, CA: Stanford University Press.  
1988 A usage-based model. In *Topics in cognitive linguistics*, B. Rudzka-Ostyn, (eds.), 127–161. Amsterdam: John Benjamins.
- Leeman, J., I. Arteagoitia, B. Fridman and C. Doughty  
1995 Integrating attention to form with meaning: Focus on form in content-based Spanish instruction. In *Attention and Awareness in Foreign Language Learning*, R. Schmidt (ed.), 217–258. University of Hawai'i at Manoa: Second Language Teaching and Curriculum Center.
- Long, M. H.  
1991 Focus on form: A design feature in language teaching methodology. In *Foreign Language Research in Cross-Cultural Perspective*, K. de Bot, R. B. Ginsberg and C. Kramersch (eds.), 39–52. Amsterdam/Philadelphia: John Benjamins Publishing Company.

- Long, M. H. and P. Robinson  
 1998 Focus on form: Theory, research, and practice. In *Focus on Form in Classroom Second Language Acquisition*, C. Doughty and J. Williams (eds.), 15–63. Cambridge, UK: Cambridge University Press.
- MacWhinney, B. and J. Leinbach  
 1991 Implementations are not conceptualizations: Revisiting the verb learning model. *Cognition* 40: 121–157.
- Marchman, V.A.  
 1997 Children’s productivity in the English past tense: The role of frequency, phonology, and neighborhood structure. *Cognitive Science* 21, 283–304.
- Markus, G. F., S. Pinker, M. Ullman, M. Hollander, T. J. Rosen and F. Xu  
 1992 *Overregularization in language acquisition*. Monographs of the Society for Research in Child Development 57 (4), Serial No. 228. Chicago: University of Chicago Press.
- Markus, G. F., U. Brinkmann, H. Clahsen, R. Weise and S. Pinker  
 1995 German inflection: The exception that proves the rule. *Cognitive Psychology* 29: 189–256.
- Matcovich, O.  
 1998 Regular inflection in the mental lexicon: Evidence from Italian. *Proceedings from the conference “The Verb in Cognitive Grammar,”* Gran.
- Orsolini, M. and W. Marslen-Wilson  
 1997 Universals in morphological representation: Evidence from Italian. *Language and Cognitive Processes* 12: 1–47.
- Orsolini, M., R. Fanari and H. Bowles  
 1998 Acquiring regular and irregular inflection in a language with verb classes. *Language and Cognitive Processes* 13: 425–464.
- Pinker, S.  
 1991 Rules of language. *Science* 253: 530–535.
- Pinker, S. and A. Prince  
 1988 On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition* 28: 73–193.  
 1991 Regular and irregular morphology and the psychological status of rules of grammar. *Proceedings of the 1991 Meeting of the Berkeley Linguistics Society*.  
 1994 Regular and irregular morphology and the psychological status of rules of grammar. In *The Reality of Linguistic Rules*, S.D. Lima, R.L. Corrigan and G.K. Iverson, (eds.), 353–388. Amsterdam: John Benjamins.
- Plunkett, K. and V. Marchman  
 1991 U-shaped learning and frequency effects in a multi-layered perception: Implications for child language acquisition. *Cognition* 38: 43–102.  
 1993 From rote learning to system building: Acquiring verb morphology in children and connectionist nets. *Cognition* 48: 21–69.

- Prasada, S. and S. Pinker  
1993 Generalization of regular and irregular morphological patterns. *Language and Cognitive Processes* 8: 1–56.
- Prasada, S., S. Pinker and W. Snyder  
1990 Some evidence that irregular forms are retrieved from memory but regular forms are rule-generated. Paper presented at *The 31<sup>st</sup> Annual Meeting of the Psychonomics Society*, 16–18 November.
- Ragnasdóttir, H., H. G. Simonsen and K. Plunkett  
1997 Acquisition of past tense inflection in Icelandic and Norwegian Children. In E. V. Clark (ed.), *Proceedings of the 28<sup>th</sup> Annual Child Language Research Forum*, Stanford.
- Robinson, P.  
1995 Aptitude, awareness and the fundamental similarity of implicit and explicit second language learning. In *Attention and Awareness in Foreign Language Learning*, R. Schmidt (ed.), 303–358. University of Hawai'i at Manoa: Second Language Teaching and Curriculum Center.  
1996 *Consciousness, Rules, and Instructed Second Language Acquisition*. New York: Peter Lang.
- Rumelhart, D. E. and J. L. McClelland  
1986 On learning the past tenses of English verbs. In *Parallel Distributed Processing: Explorations in the Microstructures of Cognition* vol. 2, J. L. McClelland and D. E. Rumelhart, (eds.), 216–271. Cambridge, MA: Bradford/MIT Press.
- Schmidt, R.  
1995a Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In *Attention and Awareness in Foreign Language Learning*, R. Schmidt (ed.), 1–63. University of Hawai'i at Manoa: Second Language Teaching and Curriculum Center.
- Schmidt, R.  
1995b Preface. In *Attention and Awareness in Foreign Language Learning*, R. Schmidt (ed.), IX–XVII. University of Hawai'i at Manoa: Second Language Teaching and Curriculum Center.
- Sharwood Smith, M.  
1993 Input enhancement in instructed SLA: Theoretical bases. *Studies in Second Language Acquisition* 15: 165–179.
- Simonsen, H. G.  
2000 Past tense acquisition and processing in Norwegian: Experimental evidence. *Language and Language Behavior* 3 (2): 86–101.
- Spada, N. and P. Lightbown  
1993 Instruction and the development of questions in L2 classrooms. *Studies in Second Language Acquisition* 15: 205–224.
- Stemberger, J. and B. MacWhinney,  
1988 Are inflected forms stored in the lexicon? In *Theoretical Morphology: Approaches in Modern Linguistics*, M. Hammond and M. Noonan, (eds.), 101–116. New York: Academic Press.

- Townsend, C. E.  
1975 *Russian Word-Formation*. Slavica Publishers.
- Trahey, M. and L. White  
1993 Positive evidence and preemption in the second language classroom. *Studies in Second Language Acquisition* 15: 181–204.
- Ullman, M. T.  
1999 Acceptability ratings of regular and irregular past-tense forms: Evidence for a dual-system model of language from word frequency and phonological neighborhood effects. *Language and Cognitive Processes* 14: 47–67.
- Ullman, M. T.  
2000 A mental model of morphology: The psychological and neural bases of the computation of complex words. *University of Maryland Working Papers in Linguistics* 10: 127–156.
- VanPatten, B.  
1996 *Input Processing and Grammar Instruction in Second Language Acquisition*. Norwood, New Jersey: Ablex Publishing Corporation.
- VanPatten, B. and T. Cadierno  
1993 Explicit instruction and input processing. *Studies in Second Language Acquisition* 15: 225–243.
- VanPatten, B. and S. Oikkenon  
1996 Explanation versus structured input in processing instruction. *Studies in Second Language Acquisition* 18: 495–510.
- White, J.  
1998 Getting the learners' attention: A typographical input enhancement study. In *Focus on Form in Classroom Second Language Acquisition*, C. Doughty and J. Williams (eds.), 85–113. Cambridge, UK: Cambridge University Press.
- Williams, J. and J. Evans  
1998 What kind of focus and on which forms? In *Focus on form in classroom second language acquisition*, C. Doughty and J. Williams (eds.), 139–155. Cambridge, UK: Cambridge University Press.
- Zalizniak, A. A.  
1980 *The Grammatical Dictionary of the Russian Language*. Moscow: Russian Language Publishers.

# **Investigating the role and effects of form-focused instruction**





## **Teaching marked linguistic structures – more about the acquisition of relative clauses by Arab learners of English**

*Ahlem Ammar and Patsy M. Lightbown*

A substantial body of second language acquisition research has revealed that learners who are taught marked relative clauses (RCs) can acquire the less marked ones without instruction on them. On the basis of this research, Hamilton (1994) proposed the Implicational Generalization Hypothesis (IGH), i.e., that knowledge of more marked forms implies the knowledge of less marked ones. In this study three experimental groups of Arabic speakers learning English as a foreign language were taught RC types with different levels of markedness. On posttest measures the experimental groups showed better command of relativization than a control group. The IGH was supported in that students generalized their knowledge of relative clauses to RC types that were implicated by the ones that were taught. However, students *also* generalized to some RC types that were more marked, and thus not implicated by the ones they were taught. This finding adds to Hamilton's questions about the unidirectionality of the IGH.

### **More about the acquisition of relative clauses by Arab learners of English**

The question of whether explicit language-focused instruction is beneficial for the second language learner has been investigated from a number of different angles. One set of studies has examined the effects of form-focused instruction on the rate and sequence of acquisition of relative clauses in a second language. Some researchers have investigated the acceleration of learning that may result from teaching relative clauses that are more "marked" according to the Noun Phrase Accessibility Hierarchy (see below). The first hypothesis shared by several studies was that instruction that focused on a marked structure might lead learners to generalize such that they would acquire not only that structure but also the less

marked structures, without specific instruction on them. Related to this was a second hypothesis, that generalization was unidirectional, that is, learners would acquire the taught structures and the less marked but not the more marked relativization types.

Considerable support for the first hypothesis has been provided by Gass (1982), Eckman, Bell and Nelson (1988), and Doughty (1991). In those studies, learners who received instruction on a marked relative clause type learnt that clause type and generalized this knowledge to the less marked types implicated by it. Findings regarding the unidirectionality hypothesis are less conclusive. Researchers found that some learners generalized new knowledge not only to the less marked but also to some more marked relative clause types. This unexpected generalization was usually considered to be either insignificant or exceptional.

In light of these findings, the present study was designed to further investigate both the effectiveness of instruction and the directionality of any generalizations.

### **The Noun Phrase Accessibility Hierarchy**

After analysing data from 50 languages, Keenan and Comrie (1977) found that noun phrases in different positions (subject (SU), direct object (DO), indirect object (IO), object of preposition (OPREP), genitive (GEN) and object of comparison (OCOMP)) are relativized according to a hierarchical ordering. The Noun Phrase Accessibility Hierarchy (NPAH) reflects the typological relationship they found among the relativizable syntactic positions of noun phrases in the languages they analyzed. According to the NPAH, the order of relativization is as follows, where “<” is read as “more accessible than” or “implicated by”.

SU < DO < IO < OPREP < GEN < OCOMP

According to Keenan and Comrie (1977) these relativization types are implicationaly related such that if a language allows relativization on a given position (e. g. OPREP), it also allows relativization on the positions to its left (in this case the IO, DO, and SU). The hierarchy is also described as an implicational scale of markedness. Keenan (1975: 138) explains that there may be some sense in which it is ‘easier’ or more ‘natural’ for relative clauses to be formed on the positions on the left than those on the right.

Gass (1982) provided the following English examples to illustrate each relativization type.

SU	The man who saw the cat...
DO	The man that the cat saw...
IO	The man that I gave the book to...
OPREP	The table that he is standing on...
GEN	The man whose book I borrowed...
OCOMP	The man that he is taller than...

### The Implicational Generalization Hypothesis

Studies of L2 relativization led to what Hamilton (1991) called the Implicational Generalization Hypothesis (IGH). Hamilton had earlier used the term “Markedness Generalization Hypothesis”, but chose to use “Implicational Generalization Hypothesis” to avoid an association with “universals or Universal Grammar, an association not necessarily warranted by the data” (Hamilton 1994: 152, footnote 1). The IGH posits two main hypotheses. First, once one relativization type has been acquired, generalization to the ones implicated by it will also take place. The generalization is *unidirectional* such that no unimplicated (more marked) relative clause types are acquired. Second, the generalization is hypothesized to be *maximal*. That is, all RC types implicated by the instructed level will be acquired.

### Research testing the Implicational Generalization Hypothesis

Gass (1982) investigated the extent to which L2 learners can acquire a less marked relativization when instruction is provided only on a more marked position. Eighteen low-intermediate ESL students whose native languages were Arabic, Italian, Russian, and Spanish were divided into one experimental group (n=13) that was taught OPREP and one control group (n=5) that was taught RCs following the order of presentation in Krohn (1977). According to Gass, “[in Krohn] the first relative clause types taught [are] subject and objects. Indirect object relative clauses are part of the category entitled objects. Genitives are introduced following these relative clause types, and with less emphasis” (Gass 1982: 132). Participants were pre-tested three days before instruction began and posttested two days after the instructional intervention ended by means of a sentence combination task and a grammaticality judgement task. The instruction was given in three half-hour sessions spread over a week and supplemented with homework

tasks. A comparison between the pretest and the posttest scores revealed that most learners (11/13) in the experimental group learned the OPREP relativization and generalized this knowledge to the less marked RCs, which were not taught. The control group's learning was limited to the RC types that were taught in class. Furthermore, some of the experimental learners generalized their learning not only to the less marked, but also to a more marked RC type (OCOMP). This finding contradicts the hypothesized unidirectionality of the IGH. Gass attributed this outcome to stranding. She explains that stranding the comparative "than" might have been interpreted by the subjects to be similar to stranding prepositions in the IO and OPREP relative clause types. However, when it came to the eight students who generalized to the unimplicated GEN by using 'who his' instead of 'whose', Gass acknowledged that these students might have generalized their learning to the GEN. She points out that errors of relative marker morphology for GEN were likely due to the fact that "the genitive marker 'whose' can not be intuited" (Gass 1982: 138).

In a replication of Gass's (1982) study, Eckman, Bell and Nelson (1988) investigated L2 learners' ability to generalize to less and more marked RC positions once one specific position was taught. Eckman, Bell and Nelson grouped 36 low-intermediate and intermediate students, who were enrolled in an ESL intensive program, in three experimental groups and one control group. Each experimental group was taught one RC type (one was taught SU, the other DO, and the third OPREP), and the control group was taught techniques of combining sentences that were not related to RCs. Instruction lasted one hour for each group, and consisted of a) a form-focused demonstration of sentence combining, b) oral exercises, and c) written exercises. Subjects' knowledge of English relativization was evaluated by means of a sentence combination task. The tests included only RCs that corresponded to the ones taught in the three experimental interventions i.e., SU, DO and OPREP. Three major findings were reported. First, participants not only learned the RC type they were taught, but also the ones implicated by it. Second, all groups did best on the structure targeted by the instruction they received. Third, the SU instructed group showed some gains with respect to the unimplicated DO. Their total error rate decreased on the posttest, but Eckman et al. did not mention whether this difference was significant. The SU group did not generalize to the OPREP type and neither did the DO group.

In an investigation of the effectiveness of meaning-oriented versus rule-oriented instruction, Doughty (1991) tested the generalization that was hypothesized to result from teaching a more marked RC types. Although

her study dealt with a variety of issues, only findings related to the generalization question will be reported here. Twenty intermediate adult ESL learners of various L1s were randomly assigned to two experimental groups and a control group. Using a time controlled computer program, Doughty was able to control two important variables, namely the amount of time and the type of instruction given to each group. The two experimental groups received instruction in ten periods spread over ten days. This was incorporated within lessons designed to build reading comprehension skills. The two experimental groups were assisted in comprehension, either by being given lexical or semantic rephrasing (the meaning-oriented group) or by being given grammatical instruction on relativization (the rule-oriented group). The RC type targeted in the instruction was OPREP. The control group was simply given more time to read.

Participants were tested by means of oral and written tasks. The latter consisted of a grammaticality judgement task devised by Doughty, a sentence-combining task adapted from Gass (1982), another sentence combining task borrowed from Ioup (1983), and a grammaticality judgement task adapted from Gass (1982). The oral measure consisted of a picture elicitation task adapted from Hyltenstam (1984). Using a 70% acquisition criterion, Doughty reports the same pattern of results that emerged from Gass (1982) and Eckman, Bell and Nelson (1988). Instruction targeting OPREP relative clauses permitted most participants to generalize to the uninstructed, but implicated, levels of the NPAH. Moreover, like some subjects from the Gass (1982) study, some subjects generalized to the unimplicated OCOMP relativization. Doughty (1988) pointed out that the mother tongues of the subjects who generalized to the unimplicated OCOMP disallowed that relativization type, thus ruling out the possibility of transfer. Doughty concluded that these subjects really did project what they learnt from OPREP instruction to the OCOMP. She summarizes the findings of previous research as well as her own:

At this point in the research, only one conclusion is clear-cut: instruction incorporating unmarked data generalizes only to unmarked contexts, whereas instruction incorporating marked data potentially generalizes not only to that marked context but to other contexts as well. How and why this happens are important topics for future research. (Doughty 1988: 52)

The inconclusive findings regarding both unidirectionality and maximality led Hamilton (1994) to further investigate the issue. In his study, Hamilton chose to use a different hierarchy, one that proposes an implicational relationship between four types of relativisation according to roles of the target

noun phrases (NPs) in both the main clause and the relative clause. In this hierarchy

the first code ... refers to the head noun as either subject (S) or direct object (O) of the matrix clause, and the second code refers to the role of the NP target of relativization within the relative clause [Only the latter had been considered in the earlier related studies.]:

- OS They saw the boy who entered the room.
- OO A man bought the clock that the woman wanted.
- SS The man who needed a job helped the woman.
- SO The dog that the woman owns bit the cat. (Hamilton 1994: 134)

Hamilton then proposed that the implicational relationship would be

OS < OO / SS < SO

where < means “implicated by”. Hamilton draws on work by O’Grady (1987) and Doughty (1988) in discussing the processing basis for this hierarchy. Basically, the more an RC type creates discontinuous processing, the more difficult it will be. Thus, for example, SO is considered most difficult because multiple discontinuities are created between syntactic units. (See also Kuno 1974; Ioup & Kruse 1977; Schumann 1980; Sheldon 1974 for discussions of this hierarchy, sometimes with different predictions about level of difficulty).

Hamilton divided 33 low-intermediate and intermediate adult ESL learners into three experimental groups (8 subjects each) and one control group (9 subjects). The three experimental groups received instruction in SO, SS, or OS relativization. The control group received instruction in sentence combining using correlative conjunctions. Each of the four groups was instructed for 45 minutes on each of 2 consecutive days.

Learners were pretested two weeks before the instruction began, and posttested two to three days after instruction ended by means of a written sentence combining task. Using an 80% acquisition criterion, Hamilton reported two main findings. Participants showed gains in almost every case on the instructed RC type, and in certain cases on those types implicated by them. More importantly, however, two OS students generalized to the more difficult types (in one case, to OO, and in the other to both SS and SO – OO having been acquired already on the pretest). This is illustrated in Table 1.

Table 1. Hamilton (1991, 1994) Pre- and post-instructional levels:  
Sentence combining task

	Learner	OS	<	OO	/	SS	<	SO
SO-instructed group	1	+		0		0		0
	2	X		+		+		0
	3	0		0		+		+
	4	X		+		+		+
	5	0		0		+		+
	6	X		X		+		+
	7	X		X		+		+
	8	X		X		+		+
SS-instructed group	9	+		0		0		0
	10	+		+		0		0
	11	X		+		0		0
	12	X		+		0		0
	13	X		X		0		0
	14	0		0		+		0
	15	X		0		X		0
	16	X		X		X		0
OS-instructed group	17	+		0		0		0
	18	+		0		0		0
	19	+		0		0		0
	20	X		0		0		0
	21	-		0		0		0
	22	+		+		0		0
	23	+		X		+		+
	24	X		X		0		0

(Adapted from Hamilton 1994: 139)

X = already acquired on pretest and maintained on posttest; + = newly acquired on posttest; - = newly lost on posttest; 0 = never acquired.

Hamilton concludes that the evidence for unidirectionality is still “somewhat ambiguous”. As for those who generalized to the more marked RCs, Hamilton argues that their exceptional performance does not alter the general pattern of the implicational generalization, which is mainly unidirectional towards the unmarked implicated positions. Nevertheless, a close look at Table 1 reveals that the number of subjects (2) who generalized to the more marked relativizations in the OS-instructed group is comparable to the number of subjects who generalized to the implicated relativization types in the SS group. In the latter, only 3 subjects generalized to the equally marked relativization type (OO), and only 2 generalized to the least marked relativization type (OS). This pattern of behaviour is very much like

the one found in Doughty (1988, 1991). Table 2 presents the performance of individual learners in Doughty (1991).

*Table 2.* Doughty (1988, 1991) Individual subjects' NPAH levels acquired

Learner	SU	<	DO	<	OPREP	<	OCOMP
11	0		0		0		0
12	+		0		0		0
14	+		0		0		0
16	+		0		0		0
1	+		+		0		0
19	+		+		+		0
4	X		0		0		0
13	X		0		0		0
15	X		+		0		0
10	X		+		+		0
6	X		+		+		0
2	X		+		+		0
7	X		+		+		+
20	X		+		+		+
5	X		+		+		+
21	X		+		+		+
17	X		X		X		+
8	X		X		X		+
9	X		X		X		+

(Adapted from Hamilton 1994 / from Doughty 1991: 451–452, 456–457)

X = Acquired on pretest; + = newly acquired on posttest; 0 = never acquired.

As shown in Table 2, 8 subjects who showed no relative clause or only SU knowledge before the instruction began acquired OPREP, the instructed RC type, and all of these generalized to the implicated DO. Furthermore, 7 subjects, 3 of whom already showed knowledge of OPREP on the pretest, also generalized to the unimplicated OCOMP.

In summary, results from Gass (1982); Eckman, Bell and Nelson (1988); Doughty (1991); and Hamilton (1994) indicate two main things. First, learners do benefit from RC instruction. Second, benefits derived from exposure to instruction on the marked types do frequently generalize to the unmarked RC types. However, the evidence regarding whether exposure to one marked type leads to generalization to more marked types, is confusing and inconclusive. All the above-mentioned studies report that some subjects were able to generalize to more marked RC types. This finding was usually treated as either insignificant or exceptional, which leaves the question of directionality open to further investigation.



## **Hypotheses**

In light of the studies that have previously demonstrated the role of instruction in accelerating the rate of acquisition of relative clauses, the first hypothesis of the present study is:

H1: Learners of English as a foreign language who receive instruction on one or all relativization types will do better than those who do not receive any instruction on relativization.

In accordance with studies that have indicated that benefits accruing from teaching one relativization type can be generalized to other relativization types, the second hypothesis of the study is:

H2: Instruction targeting marked relative clauses will generalize to unmarked contexts of relativization that are implicated by them.

Finally, taking into consideration previous research that has failed to confirm of the unidirectionality of the Implicational Generalization Hypothesis, the third hypothesis is:

H3: The generalization to other relative clause types is not strictly unidirectional. Some subjects will generalize not only to relativization types that are less marked and implicated by the instructed relative clause type of the hierarchy, but also to relativization types that are more marked and unimplicated by it.

## **Research context**

This study was carried out in a secondary school in Tunisia. All participants were native speakers of Arabic who had received most of their education in Classical Arabic with French instruction added at age 9. Before presenting the research design, it is useful to examine the differences between relativization patterns in Arabic, French and English. It should be noted that, with regard to relative clauses, there are no differences between Classical Arabic and the variety of Arabic spoken by the students.

**Differences between Arabic and English relativization systems**

Keenan and Comrie (1977, 1979), Ioup and Kruse (1977), and Schachter (1974) have investigated the differences between Arabic and English with respect to the formation of relative clauses, specifically relative marker appearance, relative marker morphology, pronoun retention, and case marking. Based on these contrastive studies, we may conclude that in general Arabic speakers encounter problems mainly with pronoun retention when learning English relative clauses. Only the findings related to pronoun retention will be reported in this study. The pattern of pronoun retention in Arabic, French and English languages is displayed in Table 3.

*Table 3.* Pronoun retention in Arabic and English

Pronoun retention	SU	DO	IO	OPREP	GEN	OCOMP
Arabic (classical)	–	+	+	+	+	+
English	–	–	–	–	–	–
French	–	–	–	–	–	n/a

(Adapted from Keenan and Comrie 1977: 93)

Arabic, unlike English, retains pronouns in all relative clause types except SU. French, like English, does not retain pronouns in any type of relativisation. Unlike English and Arabic, French does not have OCOMP relativisation. Schachter, Tyson and Diffley (1976) found that Arabic speakers who considered themselves bilingual in French were less likely to accept the grammaticality of retained pronouns than were monolingual speakers of Arabic. Examples (1) and (2) illustrate the Arabic non-retention in SU relative clauses and the obligatory retention in DO relativization respectively, as well as the corresponding English and French sentences.

- (1) Al-rajulu Ya'rifu al-fatata allati tanamu fi-tariki.  
 The man knows the girl that sleeps in the street.  
 The man knows the girl who sleeps in the street.
- L'homme connaît la fille qui dort dans la rue.  
 The man knows the girl who sleeps in the street.

- (2) Al-waladu ya'rifu ar-rajula allathi dharabat-hu as-sayaratu.  
The boy knows the man that hit him the car.  
The boy knows the man whom the car hit.
- Le garçon connaît l'homme que la voiture a frappé.  
The boy knows the man whom (that) the car hit.

### Methodology

#### Target of instruction

The present study targeted restrictive relativization. Restrictive relative clauses provide information that specifies which person or thing (the head noun) is referred to. Conversely, a non-restrictive relative clause provides further information that is not needed to identify the head noun referred to. Examples (3) and (4) illustrate a restrictive relative clause and a non-restrictive one respectively.

- (3) He was talking to the girl who was sitting in front of him.  
(4) Nadia, who is sitting in front of John, is Tunisian.

The instructional intervention was further limited in that only the contrasting rules for pronoun retention in Arabic and English were explained and tested.

#### Participants

The primary criterion in selecting participants was that they had very little or no knowledge of relativization. A grammaticality judgement task to measure participants' knowledge of relativization was administered as a pretest three days prior to the experimental instruction. To be included in the study students had to demonstrate the ability to correctly judge at least two distracters (sentences including errors not related to relative clauses). Students who were able to correct half or more of the relativization sentences that they judged as incorrect were excluded as already too advanced. Only those students who correctly judged the grammaticality of fewer than half the relative clause items but could not correct them and those who could neither judge nor correct relative clauses were selected for the study.

The group of students who were eventually included in the study were 34 low-intermediate Tunisian students all studying English as a foreign

language for the second year in their secondary education. They were drawn from two classes but all had the same teacher for their regular English classes. The instruction in those classes included a range of activities – both communicative and grammar-focussed lessons (see Ammar, 1996). The participants were divided into four groups, three experimental and one control group. Assigning the participants to the different groups was not random due to practical factors beyond our control.

One experimental group, referred to as DOG, was taught the DO relativization. The second group, referred to as OPG, was taught OPREP. The third was exposed to SU, DO and OPREP relative clauses and is referred to as SDO. The control group, COG, was taught how to combine sentences using the conjunctions “while”, “before”, and “after”.

The first experimental group was taught the DO relative clause type to better test the directionality question. In previous studies which tested the IGH (Doughty 1991; Eckman, Bell and Nelson 1988; Gass 1982) the major exception to the unidirectionality of the IGH was found with the OCOMP relative clause type. Participants showed more success in generalizing to the OCOMP than to the GEN when taught the OPREP. Gass (1982) attributed this to a possible confusion that students might have had between prepositions in IO and OPREP relative clause types and the comparative “than” in OCOMP. The inclusion of a group that is not taught any relative clause type containing prepositions (DOG) may help to determine whether the confusion is the basis for generalization to the unimplicated OCOMP. It also provides more opportunities to test the directionality of generalization. That is, contrary to some previous studies (Doughty 1991; Gass 1982) which taught the OPREP and which, consequently, could test the generalization to more marked RCs in only two positions (GEN and OCOMP), the DOG in this study permitted two more opportunities to test the directionality of that generalization. In fact, if generalization to more marked positions is possible, the DOG participants in this study would have four chances (IO, OPREP, GEN and OCOMP) to do so.

### Pretesting

The control and experimental groups were pretested using a grammaticality judgement task that was adapted from Gass (1982) and Doughty (1991) (see Appendix 1). Participants were given 4 sentences for each of the six relative clause types on the NPAH. Two of the sentences contained pronoun retention errors and the other two were correct. Examples 5 and 6 show a correct and an incorrect Direct Object relative clause sentence respectively.

(5) The book that I read yesterday is very interesting.

(6) \*The test that I gave it was very difficult.

Apart from the 24 sentences that were included to test the participants' knowledge of English relativization, the test included eight distracters that contained pluralization and subject verb agreement errors. These two grammatical structures were chosen because they occur very early in the participants' program and tend to be reviewed frequently. Thus one would anticipate that students who were focused enough and who understood how to perform the grammaticality judgement would have high accuracy rates on these items. Sentences (7) and (8) illustrate the two types of distracters.

(7) \*The students is going on a journey to Monastir next Sunday.

(8) \*Tunisia has many beautiful beach and hotels.

Even though the lexicon was simple, students were encouraged to ask about the meaning of any unfamiliar vocabulary item in the test. The 32 sentences were presented in an order that was intended to ensure that participants did not view several consecutive examples of the same RC type that differed only by the retained pronoun. Students were instructed to read each sentence and decide whether it was a grammatical or an ungrammatical English sentence. If the sentence was correct, students were asked to circle "yes", if not they had to circle "no". To avoid having learners focus on spelling, the participants were told that there were no spelling mistakes in the target items. Once they finished judging all the sentences, students were asked to correct the sentences they had judged as incorrect. Participants did not know that they would eventually have to do the corrections before they began. This was done to avoid the possibility that students would avoid judging sentences as incorrect in order to avoid the additional challenge of making corrections.

### Posttesting

One day after the instructional intervention ended, the experimental groups and the control group were tested using two tasks: a grammaticality judgement task and a sentence combination task. In the grammaticality judgement task, participants were required to judge 48 sentences (8 sentences for each of the 6 RC types). Four sentences of each type were correct and

four contained pronoun retention errors. No distracters were included. Participants were instructed to perform the same task as they did in the pretest. At this point, students from the experimental groups would certainly be expected to know what was being targeted.

The sentence combination task was adapted from Gass (1982) and Doughty (1991). Participants were asked to combine 12 pairs of simple sentences into single sentences containing the information from both original sentences. Participants were instructed to begin with the first simple sentence and not to use the conjunctions “because”, “while”, “after”, “since”, “before”, or “and”. The randomly ordered sentences included two tokens of each of the six relative clause types when correctly combined. Example (9) shows an original pair together with the resulting sentence containing a subject relative clause.

- (9) The book is new. The book is in the bag.  
The book that is in the bag is new.

Given that the control group participants had not been taught relative clauses and that the instruction to which they were exposed targeted combining sentences using “while”, “before”, and “after”, the instructions did not warn the COG against using any of these conjunctions. The instructions were rather as follows:

Please combine the two sentences into one correct English sentence.  
Always begin with the first sentence.  
Do not leave out any information.

These changes were adapted to minimize the extent of frustration that might be felt by the COG participants. Two weeks after the first posttest the experimental groups’ knowledge of RCs was tested again using the same tasks.

#### Instructional intervention

The instruction presented to the four groups consisted of two main parts: an in-class part consisting of three thirty-minute sessions for each group and a take-home part adapted from “homework packets” developed by Croteau (1995).

The experimental groups and the control group were exposed to the same types of in-class activities. All instruction was carried out by the first

author in a classroom made available by the school. Students came after school hours to participate in the three 30-minute in-class sessions. They consisted of (1) an elicitation activity, (2) a rule presentation activity and (3) practice activities.

In the elicitation activity, students were asked to describe some people and actions in four different pictures. The rule presentation activity was a teacher-led version of Doughty's (1991) sentence combination procedure. Starting with some of the sentences from the elicitation activity, the teacher (experimenter) explained the procedure for combining two simple sentences into one complex sentence containing a relative clause. The procedure involved the following steps: the identification of the co-referential noun; information about turning the co-referential noun into a relative pronoun; the movement of the relative pronoun to the correct position; the explanation of the deletion of the co-referential pronoun; and finally, the placement of the relative clause in the target position. All these steps were accompanied by an explicit statement of the relevant rules.

The rule presentation activity was experimentally manipulated so that each experimental group viewed the targeted relative clause type only. For instance, the DOG viewed sentences that could be combined into complex sentences containing DO relative clauses only. After the sentence combination procedure had been explained, participants were exposed to the pronoun retention problem. Students were given examples of relative clauses with retained pronouns and told that the co-referential noun should be replaced only once i.e., with a relative pronoun. A comparison of the Arabic and English relativization systems was provided to further explain the pronoun retention problem. The rule presentation activity for the COG consisted of an explanation of the sentence combination procedure using the conjunctions "while", "before" and "after".

Three types of practice activities were used: grammaticality judgement, scrambled sentences and sentence combination. The instructions for the grammaticality judgement and the sentence combination were identical to the ones given in the pre- and posttests. In the scrambled sentence exercises, the sentences were divided into blocks of two sentences and participants were instructed to put the words in the right order to obtain sentences with relative clauses (for the experimental groups) or with "while", "after" and "before" (for the COG). Each exercise given in the practice activity contained 8 sentences.

The "homework packets" were structured in a way that paralleled the content of the in-class instruction. Both rule presentation and practice parts were included. In the former, the participants were presented with a step-

by-step explanation of the sentence combination procedure as it occurred in the in-class rule presentation activity. This part was manipulated such that each group viewed the targeted structure only i.e., DO for the DOG, OPREP for the OPG, SU, DO and OPREP for the SDO, and “while”, “after” and “before” for the COG. The practice part comprised the same exercises given in class i.e., grammaticality judgement, scrambled sentences, and sentence combination.

## Results

### Pretest knowledge of relativization

An analysis of variance (ANOVA) was conducted to compare the relativization ability of the four groups at the outset of the experiment. Results indicated no significant difference between the groups. See Table 4.

*Table 4.* ANOVA comparing groups' pre-instructional performance on relativization

Group	N	M	SD
DOG	7	12.71	2.87
OPG	9	12.56	2.55
SDO	8	11.38	3.62
COG	9	12.33	2.87

$F(3,29) = 0.32, P = .81$

Note: The maximum score on the pretest = 36 points: 12 for correct sentences, 12 for incorrect ones, and 12 for the expected corrections.

The pretest results were analyzed in terms of response patterns for the different RC types. The correct and incorrect sentences were examined separately. Participants in the four groups were biased towards accepting all the sentences regardless of whether they were correct or incorrect (see Tables 5, 6, and 7).

The students' overall behaviour may have two possible interpretations. The bias might be evidence of their limited RC knowledge. Alternatively, a lack of concentration on the part of the participants on the day of the test could be an explanation for this performance. That is, instead of making an



Table 5. Pretest accuracy on SU, DO, IO

Group	SUB			DO			IO		
	Gram. %	Ungram. %	COR %	Gram. %	Ungram. %	COR %	Gram. %	Ungram. %	COR %
DOG	79	14	0	86	21	0	57	35	0
OPG	72	39	14	89	22	25	28	67	0
SDO	81	25	50	75	19	0	44	56	0
COG	72	22	0	67	39	14	39	78	0

Gram. = accuracy on correct sentences; Ungram. = accuracy on incorrect sentences; COR = accuracy on correction.

Table 6. Pretest accuracy on OPREP, GEN and OCOMP

Group	OPREP			GEN			OCOMP		
	Gram. %	Ungram. %	COR %	Gram. %	Ungram. %	COR %	Gram. %	Ungram. %	COR %
DOG	86	14	0	50	29	0	86	36	0
OPG	63	22	0	56	44	25	61	17	33
SDO	75	13	0	38	19	33	75	0	0
COG	48	50	0	33	33	0	78	56	0

Gram. = accuracy on correct sentences; Ungram = accuracy on incorrect sentences; COR = accuracy on correction.

Table 7. Overall response pattern on the pretest relative clause sentences

Group	Gram. %	Ungram. %	Unjudged %
DOG	72	21	7
OPG	57	33	10
SDO	66	23	11
COG	55	43	3

Gram = judged grammatical; Ungram = judged ungrammatical; Unjudged = items left blank

effort to find the difference between the correct and the incorrect sentences, students may have just circled “yes” in the majority of the cases in order to get through the task they were required to do. This interpretation is less plausible, however, when we look more closely at their performance on the RC items. First, the tendency to accept most sentences was particularly strong with SU, DO, OPREP, and OCOMP relative clauses. With the IO relative clauses, they tended to reject instead of accept all sentences. With the GEN, they accepted the grammatical sentences and rejected the ungrammatical ones to a large extent. However, even when they rejected ungrammatical sentences, no one in the four groups ever supplied the correct form for GEN.

Further evidence that students were making intentional choices rather than randomly accepting and rejecting sentences comes from their performance on the distracter items. Students showed a different behaviour while judging these sentences. As shown in Table 8, participants in the four groups succeeded in rejecting nearly half of the ungrammatical distracters. Furthermore, when they rejected a sentence, they usually supplied the correct alternative. The students’ relative success in identifying the incorrect distracter sentences confirms that they were making an effort to do the task and that their poor performance on relative clause sentences was mainly due to their lack of knowledge of that structure. One student (in the DOG) who showed a strong bias towards accepting all the relative clause and distracter sentences was excluded from the experiment. This reduced the number of DOG participants to 7. This student’s complete bias was interpreted to be a sign of either an absence of any serious consideration of the task or an overall ability level that was so low that the participant could not benefit from the instruction.

*Table 8.* Response pattern on distracters (ungrammatical sentences only)

Group	Gram. %	Ungram. %	Unjudged %
DOG	48	48	4
OPG	39	56	6
SDO	14	75	11
COG	54	44	1

Gram. = judged grammatical; Ungram = judged ungrammatical; Unjudged = items left blank.

## Post-instructional knowledge of relativization

## Hypothesis 1:

Students who receive instruction on one or all relativization types will do better than those who do not receive any instruction on relativization.

Table 9. ANCOVA comparing groups' relativization knowledge on the posttest

Group	N	M	SD
DOG	7	68.14	4.53
OPG	9	64.78	11.85
SDO	8	66.25	10.08
COG	9	19.89	10.81

$F(3,28) = 46.95, P < .001$

Note: The maximum score on the posttest = 84 points: 24 for correct sentences, 24 for incorrect sentences, 24 for the expected corrections on the judgement task; 12 for the sentence combination task.

This hypothesis was confirmed. An analysis of co-variance (ANCOVA) using the pretest score as a co-variate revealed that the posttest difference between the groups was significant (see Table 9). Tukey analyses revealed that the three experimental groups significantly outperformed the control group ( $p < .001$ ). However, there were no differences between pairs of experimental groups. Separate ANCOVAs were run on the four groups' performance on each RC type. For each type, the pattern was the same. There were no significant differences between the three experimental groups, and each one of them did significantly better than the control group on all RC types except for the GEN. Overall, both the ANCOVA run on the whole performance, and the separate ANCOVAs run on different relative clause types showed that the experimental groups performed significantly better than the control group in relativization knowledge, which supports the first hypothesis. The significant improvement in RC knowledge from the pretest to the first posttest was maintained or increased by the three experimental groups on the second posttest. The stability of the improvement is represented in Figure 1.

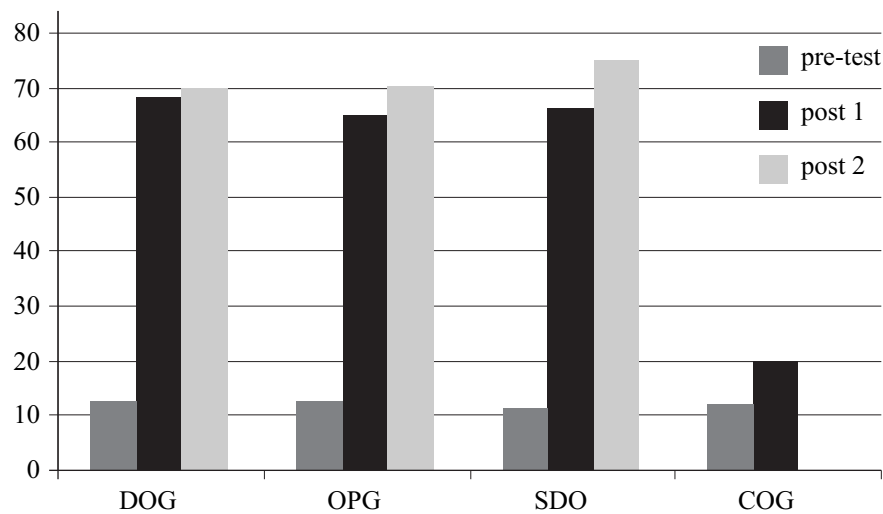


Figure 1.

#### Hypothesis 2:

Instruction targeting marked relative clauses will generalize to unmarked contexts of relativization that are implicated by them.

Overall, this hypothesis was also supported. The performance of individual students was analysed. Data were converted to a binary form representing whether or not a student had “acquired” a particular RC type on the pretest or the posttest. To be treated as having acquired a relative clause type in the pretest, a student had to correctly judge all four sentences for that relativization type (2 correct and 2 incorrect) and supply a grammatical alternative for at least one of the two ungrammatical sentences i. e., a participant had to get a score of 5 out of 6, 4 for the correct judgements and 1 for the correction. On the posttest, a participant was considered to have acquired a relativization type if s/he obtained an overall score of 12 out of 14 on that type: 1 out of 2 for the sentence combination task, and 11 out of twelve for the grammaticality judgement task (8 for the correct judgements of the grammatical and the ungrammatical sentences and 3 for supplying 3 out 4 grammatical corrections). The results are represented in Table 10.

It is evident from Table 10 that participants in the three experimental groups learned more than the instructed relative clause type(s). Some participants were able to generalize the relativization knowledge they internalized from the instruction to all implicated and less marked relativization types.

Table 10. Implicational scale for knowledge of relativization on the first posttest (based on Grammaticality Judgement and Sentence Combining Tasks)

Student	SU	DO	IO	OPREP	GEN	OCOMP
<b>DOG</b>						
1	+	+	0	0	0	+
2	+	+	+	+	0	+
3	+	+	+	0	0	+
4	0	+	+	0	0	+
5	+	+	0	0	0	+
6	+	+	+	+	0	+
7	+	+	+	+	0	+
<b>OPG</b>						
8	0	+	0	0	0	+
9	0	0	0	0	0	+
10	+	+	+	+	0	+
11	0	+	+	+	0	+
12	+	+	+	+	+	+
13	+	+	0	+	0	+
14	+	+	+	+	0	+
15	0	0	+	+	0	+
16	0	+	+	+	0	+
<b>SDO</b>						
17	+	+	0	+	0	+
18	+	+	+	+	0	+
19	+	+	+	+	0	+
20	+	0	0	0	0	-
21	+	+	+	+	0	+
22	+	+	+	+	0	+
23	+	+	+	+	0	+
24	+	0	0	0	0	0

+ = newly acquired on posttest; 0 = never acquired. (Adapted from Ammar 1996)

In the DOG, six participants out of seven were able to generalize to the implicated and untaught SU. Three (#10, #12 and #14 ) of the OPG participants were able to generalize to all three implicated relative clause types. The remaining 6 participants generalized the relativization knowledge acquired from the instruction on OPREP to one or two implicated positions of the NPAH.

#### Hypothesis 3:

The generalization to other relative clause types is not strictly unidirectional. Some participants will generalize not only to relativization types that are less marked and implicated by the instructed relative clause level, but also to relativization types that are more marked and unimplicated by it.

This hypothesis was supported. Table 10 shows that some students in the DOG and OPG generalized not only to the RC types implicated by the relativization targeted by the instruction, but also to the relative clause types that were not implicated by it. Three participants (#2, #6 and #7) in the DOG were able to generalize to all the unimplicated more marked relative clause types except for the GEN. The remaining 4 participants generalized to at least one more marked level of the NPAH. Table 10 shows that the GEN was the only relative clause type to which no generalization took place. The same pattern of generalization was found with the OPG and SDO. Nearly all students generalized to the most marked relative clause type i.e., OCOMP. At the same time no one was able to perform successfully on the items exemplifying GEN. The generalization to the more marked relative clauses runs counter to the hypothesised unidirectionality of the IGH and supports the hypothesis that there would be bi-directional generalization.

### **Discussion**

Three major findings emerge from this study. First, the experimental findings suggest that exposure to and instruction about relative clauses have strong positive effects. While the control group, which was exposed to structures other than relative clauses, did not show any gains in relativization ability, the experimental groups' knowledge of that structure improved significantly. In fact, the students benefited equally well from instruction targeting one (DOG and OPG) or more than one RC type (SDO). Second, the study adds further support to the hypothesized generalization that is

expected to take place once a marked relative clause type is taught. Participants in the OPG and DOG were able to project the rules they internalized from the instructed relative clause type to the implicated and less marked contexts of relativization which were not targeted by the instruction. Finally, the results add to the evidence that generalization is not unidirectional. Nearly all learners from the three experimental groups were able to project the rules of the relativization system not only to the less marked contexts but also to the more marked ones, with the apparent exception of the genitive.

### **Generalization**

From a first look at the pattern of the generalization that occurred in the OPG and SDO (participants generalized to the OCOMP), the only interpretation that seems logical is the one provided by Gass (1982). Gass attributed this generalization to the students' inability to differentiate syntactically between the comparative conjunction "than" and prepositions in the OPREP. This interpretation becomes more plausible when we consider the participants' inability to generalize to the GEN that is less marked than the OCOMP. It is noteworthy, however, that even though participants in the DOG were not taught relative clauses containing prepositions (IO or OPREP), some of them generalized to the unimplicated OCOMP. Furthermore, two students managed to project their relativization knowledge to the OCOMP without generalizing to the less marked IO and OPREP relative clause types that contain prepositions. These instances suggest that the students were not treating the comparative conjunction "than" as a preposition because they apparently had less difficulty relativizing the OCOMP than IO and OPREP. Nor does it seem plausible that this generalization can be attributed to students' knowledge of French, their second language. French does not allow relativization of the OCOMP.

As suggested by Doughty, the overall unexpected generalization that occurred to more than one more marked relative clause type in the experimental groups might be attributed to the participants' generalization of the relativization system that they have deduced from the instruction provided on one relatively marked position (DO for DOG and OPREP for OPG). In the case of the present study, the participants' mother tongue might have provided further help to understand that system and to apply it to all the remaining relativizations regardless of their nature i.e., marked or unmarked. In fact, except for the pronoun retention problem that differentiates the Arabic and English relativization systems, both languages respect nearly

the same rules when it comes to RC formation. The similarity between the two systems may have reinforced the participants' hypotheses about English relative clauses. This interpretation becomes more plausible when we consider the students' inability to generalize to the GEN. Seeing that the GEN formation in English requires a morphological form that is quite different from the ones used in the other RC types targeted by the instructional intervention, the participants were neither able to judge the GEN relativization sentences nor provide the correct alternatives for that same RC type. The participants' possible reliance on the similarity existing between English and Arabic in the formation of the other relative clauses might have contributed to their poor performance on the GEN sentences. As explained by Kharma and Hajjaj, Arabic does not have an equivalent to the English pronoun "whose".

...instead in such cases Arabic employs the same relative pronoun used in the subjective or objective case depending on the case of the noun preceding it (i. e. the antecedent), whereas the genitive case is exhibited later affecting the subject or object of the relative clause.

(Kharma & Hajjaj 1989: 127)

The participants' knowledge of French might also have affected their acquisition of the English relative clause system. French is similar to English in that it does not retain pronouns but it is unlikely that it was helpful for the genitive. The genitive in French requires a unique relative marker (*dont*), one which is a late acquisition for many second language learners and which is actually not used in some varieties of French. The standard French version shown in example (10) is often rendered in the non-standard version shown in (11). This suggests that French would not be a ready source of hypotheses regarding how the genitive works.

(10) La fille dont le père travaille à l'école...  
The girl of whom the father works at the school...  
The girl whose father works at the school...

(11) La fille que le père travaille à l'école...  
The girl that the father works at the school...

In conclusion, if we assume that the participants developed one system that they used with all relative clause types, and if we further assume that the similarity between Arabic and English relativization systems, on the one hand, and between English and French, on the other hand, facilitated the



development and the reinforcement of that system, this may explain why some students generalized to the unimplicated relative clause types except for the GEN.

### **Effects of instruction**

The performance of the experimental groups when compared with that of the control group may be explained by several factors, including the highly explicit nature of the instructional treatments, the multiple levels of exposure, and the focus on a single error type in the teaching and on the pretest/posttest tasks.

#### *Explicitness.*

Rutherford and Sharwood Smith describe a range of instructional approaches that vary in terms of explicitness:

One can explicitly call attention to a grammatical feature and, if necessary, even articulate an informal pedagogical 'rule' as an instructional aid; one can implicitly call attention to a grammatical feature through calculated exposure of the learner to crucial pre-selected data; and one can choose to ignore a grammatical feature altogether.

(Rutherford & Sharwood Smith 1988: 277)

The experimental instruction provided in the present study is clearly of the explicit type. Participants were not only exposed to the different steps that enabled them to combine two simple sentences into one single sentence containing a relative clause, they were also explicitly provided with the rules behind these different steps. As explained by Doughty (1988, 1991), during the sentence combination process that schematised the transformations that the two simple sentences underwent, participants were able to observe how one co-referential noun turned into a relative pronoun, see the relative pronoun move to the correct position, notice how the co-referential pronoun was deleted, and finally see how the relative clause itself moved to the target position. The same applies to the deletion of the redundant pronoun. All the transformational steps were accompanied by an explicit statement of the rules. For instance, participants were given the rules for choosing the appropriate relative pronoun and a comparison between the Arabic and the English rules was drawn to explain the pronoun retention problem that Arab learners of English often experience. Learners had every opportunity to know exactly what the focus of the instruction was. In

addition, the fact that the lessons were taught by a visitor (the researcher/first author) would have increased the saliency of the target structure by setting it apart from regular lessons.

*Levels of Exposure.*

The instruction was designed to provide the experimental participants with multiple levels of exposure to relative clauses. First, the pretest provided all participants with a non-negligible opportunity to notice the way relative clauses were formed. Second, the instructional intervention provided a significant amount of exposure. The instruction, as previously described, started (in the first instructional period) with an overt explanation of the procedure underlying relative clause formation. The explanation of this procedure was provided at the beginning of each of the two remaining instructional periods, each time participants in any experimental group showed incomprehension of some aspects of relativization, during the explanation of the pronoun retention problem, and in the take-home materials. Finally, the exercises (in-class and take-home) not only enabled the participants to practice the recombination procedure, but also provided them with many correctly formed examples of the relative clauses on which they received instruction (DO or OPREP, or SU, DO, and OPREP).

Given these multiple levels of exposure one might argue that explicitness was not the crucial factor in the post-instructional relativization knowledge and that exposure alone explains students' gains in this study. It is impossible to refute this suggestion completely. However, the students' performance on the GEN and OCOMP relative clauses suggests the limitations of exposure alone. The analyses showed that while nearly all the experimental participants generalized to the OCOMP, no group generalized to the GEN even though, according to the NPAH, it is less marked than the OCOMP. If exposure alone were enough to lead the experimental participants to generalize to more marked positions on the posttests, generalization to the GEN should have occurred. Since it did not, we may conclude that these participants needed some explicit instruction about the different morphological forms required for the formation of the GEN. The need for this explicit explanation becomes particularly evident when these participants' L1 is taken into account. Arabic forms the GEN by following the same procedure involved in the formation of all the other RC types. The relative pronoun "whose" does not have an equivalent in Arabic. And, as noted above, French does not offer an easily transferable pattern. There were correct examples of "whose" in the test materials but students didn't pick up on it. There were, however, also examples of the incorrect forms "who his" and "who her", and students may

have taken these as confirmation of their interlanguage form. As previous research has shown, L2 learners may perceive correct forms as alternatives to their interlanguage forms, rather than as necessary replacements for them (White 1991).

*Error type.*

This study was designed to investigate Arab learners' ability to overcome pronoun retention when learning relative clauses. All other relativization strategies, namely, relative pronoun appearance, relative pronoun morphology and case, were excluded from the analysis of learners' performance. This implies that the participants' task in the grammaticality judgement tasks was limited to detecting those pronoun retention problems and correcting them. Furthermore, participants were not penalized for the wrong choice or the incorrect omission and retention of relative pronouns in the sentence combination task. This means that the participants who were considered as having acquired certain relative clause types would not be treated as such if the study had been designed to account for more than the pronoun retention strategy. An investigation of the same participants' ability to respect all the rules pertaining to English relativization might yield different results. Proponents of the unidirectionality of the IGH might argue that if participants were tested on all aspects of relativization i.e., pronoun retention, relative pronoun morphology and case, and relative pronoun appearance, no generalization to the unimplicated relative clause positions would take place. This argument is plausible but the same results might be obtained if participants were provided with 1) ample exposure to RCs and 2) with instruction that explicitly explains the pedagogical rules that apply to each of these aspects of English relative clauses.

## **Conclusion**

Three pedagogical conclusions can be drawn from these findings.

First, instruction focusing Tunisian Arabic-speaking students' attention on the problem of pronoun retention seemed to facilitate the comprehension and acquisition of the English relativization system. This instruction helped them see the major difference that exists between the Arabic and the English relativization systems, and, subsequently, enabled them to avoid the error (pronoun retention) that many Arabic speakers frequently make. This was true at least in contexts where learners were given time to think about form as is the case with the paper-and-pencil-tests administered in this study.

Missing from this study is a measure of students' ability to use various types of relative clauses in spontaneous speech and writing.

Second, this study has provided further confirmation of the hypothesis that teaching only one marked relative clause type is as beneficial as teaching several relative clause types, beginning with the least marked types. One might think that in the long term (more than the two weeks of the present study) instruction covering many contexts would be more effective. However, until another study investigates the long-term effects of teaching many relative clause types as opposed to teaching only one marked type, the evidence from this study suggests that there is no such benefit. A larger number of subjects needs to be included in future studies in order to have sufficient data to better test this hypothesis.

Third, this study clearly indicates that students cannot generalize to the correct form of the genitive no matter what RC types are taught. Instruction, targeting RC types other than the GEN, does not provide learners with any clue about the relative pronoun "whose". This implies that Tunisian learners, and Arabic-speaking learners of English in general, need more guidance in perceiving the pattern in the formation of the genitive. Because the genitive clause in English requires a different relative pronoun and a different procedure, an explanation of the use of "whose" as well as the operations involved in forming genitive relative clauses is required. Teaching one relative clause type and expecting students to generalize to the genitive did not work in this study and is not likely to be effective, even if the pronoun retention problem is explained.

Overall, the findings of this study in the context of Arabic speakers learning English as a foreign language are consistent with those of previous research with speakers of Arabic and other languages learning English in the context of English as a second language. Nevertheless, further research is needed to determine the impact of instructional intervention on learners' spontaneous production (especially oral production) of relative clauses and the long-term effects of the instruction.

### **Note**

This research was carried out for the first author's M.A. thesis at Concordia University, under the supervision of the second author. Financial support for the research came from grants to P. M. Lightbown and N. Spada from the Social Science and Humanities Research Council of Canada and the *Fonds pour la Formation de chercheurs et l'aide à la recherche* from the Government of Quebec. We thank Nina Spada for comments on an earlier version of this paper.



15. The girl who I talked about is very smart.	yes	no
16. My brother can eat two sandwich in five minutes.	yes	no
17. The boy who he stole the sandwich was caught by the policeman.	yes	no
18. I know the man who Mary is older than.	yes	no
19. The new dress that my mother bought is very beautiful.	yes	no
20. The students is going on a journey to Monastir next Sunday.	yes	no
21. The boy who his leg was broken went to the hospital.	yes	no
22. The student who the teacher gave the homework to was absent.	yes	no
23. The cat that it chased the mouse is very big.	yes	no
24. The teacher gave us many exercise in the test.	yes	no
25. The book that I'm looking for is blue.	yes	no
26. The father whose son had an accident is very sad.	yes	no
27. The boy who I am taller than is Ali.	yes	no
28. My brother play basket-ball.	yes	no
29. The girl who her father died is my classmate.	yes	no
30. The sandwich that my mother prepared it is delicious.	yes	no
31. She know how to play chess.	yes	no
32. The boy who I am older than him is Karim.	yes	no

Now, try to write your correction underneath each incorrect sentence.

EXAMPLE: The boy play hand-ball.      yes      no  
*The boy plays hand-ball.*

## References

- Ammar, A.  
1996 Is implicational generalization unidirectional and applicable in foreign contexts? Evidence from relativization instruction in a foreign language. M. A. diss., Concordia University: Montreal.
- Croteau, K. C.  
1995 Second language acquisition of relative clause structures by learners of Italian. In *Second language acquisition theory and pedagogy*, F. R. Eckman, D. Highland, P. W. Lee, J. Mileham and R. R. Weber (eds.), 115–28. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Doughty, C.  
1988 The effect of instruction on the acquisition of relativization in English as a second language. Ph.D. diss., University of Pennsylvania, Philadelphia.  
1991 Second language instruction does make a difference: evidence from an empirical study of SL relativization. *Studies in Second Language Acquisition* 13: 431–69.
- Eckman, F. R., L. Bell and D. Nelson  
1988 On the generalization of relative clause instruction in the acquisition of English as a second language. *Applied Linguistics* 9: 1–20.
- Gass, S. M.  
1982 From theory to practice. In *On TESOL '81: selected papers from the Fifteenth Annual Conference of Teachers of English to Speakers of Other Languages*, M. Hines and W. Rutherford, (eds.), 129–39. Washington DC: TESOL.
- Hamilton, R.  
1991 The markedness generalization hypothesis and the teachability hypothesis: Instructional effects on adult ESL learners' knowledge of restrictive relative clauses. M.A. diss., University of South Carolina, Columbia.  
1994 Is implicational generalization unidirectional and maximal? Evidence from relativisation instruction in a second language. *Language Learning* 44: 123–57.
- Hyltenstam, K.  
1984 The use of typological markedness conditions as predictors in second language acquisition: the case of pronominal copies in relative clauses. In *Second languages: a crosslinguistic perspective*, R. Andersen, (ed.), 39–60. Rowley, MA: Newbury House.
- Ioup, G.  
1983 Acquiring complex sentences in English. In *Second language acquisition studies*, K. Bailey, M. Long and S. Peck, (eds.), 25–40. Rowley, MA: Newbury House.

- Ioup, G. and A. Kruse  
 1977 Interference vs. structural complexity in second language acquisition: Language universals as a basis for natural sequencing. In *On TESOL '77 – teaching and learning English as a second language: trends in research and practice*, H. Brown, C. Yorio and R. Crymes, (eds.), 159–71. Washington, DC: TESOL.
- Keenan, E.  
 1975 Variation in universal grammar. In *Analyzing variation in language*, R. Fasold and R. Shuy, (eds.), 136–48. Washington, DC: Georgetown University Press.
- Keenan, E. and B. Comrie  
 1977 Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8: 63–99.  
 1979 Data on the noun phrase accessibility hierarchy. *Language* 55: 333–51.
- Kharma, N. and A. Hajjaj  
 1989 *Errors in English among Arabic speakers: Analysis and remedy*. Essex, England: Longman.
- Krohn, R.  
 1978 *English sentence structure*. Ann Arbor: University of Michigan Press.
- Kuno, S.  
 1974 The position of relative clauses and conjunctions. *Linguistic Inquiry* 5, 117–136.
- O'Grady, W.  
 1987 *Principles of grammar learning*. Chicago, IL: University of Chicago Press.
- Rutherford, W. and M. Sharwood Smith (eds.)  
 1988. *Grammar and second language teaching*. New York: Newbury House.
- Schachter, J.  
 1974 An error in error analysis. *Language Learning* 27: 205–14.
- Schachter, J., A. Tyson and F. Diffley  
 1976 Learner intuitions of grammaticality. *Language Learning* 26: 67–76.
- Schumann, J.  
 1980 The acquisition of English relative clauses by second language learners. In *Research in second language acquisition*, R. Scarcella and S. Krashen (eds.), Rowley, MA: Newbury House.
- Sheldon, A.  
 1974 The role of parallel function in the acquisition of relative clauses in English. *Journal of Verbal Learning and Verbal Behavior* 13: 272–81.
- White, L.  
 1991 Adverb placement in second language acquisition: Some effects of positive and negative evidence in the classroom. *Second Language Research* 7 : 133–161.



## **The importance of form/meaning mappings in explicit form-focused instruction**

*Nina Spada<sup>1</sup>, Patsy M. Lightbown, and Joanna L. White*

Research in second language acquisition confirms that the linguistic features learners are most likely to have long-term difficulty acquiring are those in which there is a misleading similarity between the L1 and the L2. This difficulty may be particularly great in classrooms where learners share the same L1, resulting in successful communication even when students are making errors on features influenced by their L1. Previous research shows that form-focused instruction can help learners make progress toward correct use of these features. We have hypothesized that it may be necessary to provide instruction that is not only explicit with regard to the L2, but that also draws attention to the specific differences between L1 and L2 (Spada & Lightbown 1999).

Participants in this study were approximately 90 French-speaking students (11–12 years old) in four intact classes. They were in the second half of a five-month communicative intensive ESL course. Two classes received instruction on the possessive determiners his and her, and two others received instruction on question formation (with a focus on inversion with noun subjects). The instruction included information directing the students' attention to the contrasts between their L1, French, and their L2, English. Over four weeks, students engaged in a variety of activities that gave them opportunities to hear and produce the target features. They were given a "rule of thumb" for either possessive determiners or question inversion. Throughout the instructional period, they were reminded of the rule of thumb and given corrective feedback by their teachers when they made errors on the feature that was in focus in their class. The question classes and possessive determiner classes served as uninstructed comparison groups for each other. Pretests and posttests consisted of oral production and paper and pencil tasks. Following the completion of the posttests, students were interviewed to explore their metalinguistic awareness of possessive determiners.

The classes that received instruction on possessive determiners outperformed the comparison classes on tasks assessing their knowledge of this feature. The interview data showed evidence of learning in the instructed

group, even among students whose performance on the oral production and paper and pencil tasks did not suggest much improvement. The classes that received instruction on questions also improved on the feature they were taught, but the changes were less substantial and, on most measures, not greater than those in the uninstructed group.

One explanation for the findings is related to differences between the linguistic features under investigation. A misused or misunderstood possessive determiner is more likely to cause communication problems than a question without inversion. Thus, the greater effect of instruction on possessive determiners points to the influence of form/meaning mappings on the effectiveness of explicit form-focused instruction.

### **Introduction**

The pedagogy of communicative language teaching (CLT) has come to characterize second language classrooms in a number of places throughout the world. Many teachers and second language acquisition (SLA) researchers believe that learners who are engaged in second language (L2) interaction with a focus on meaning can, at the same time, progress in their knowledge of the vocabulary, syntax, and morphology of the L2. This “two for one” approach – the ability to acquire language while focused on meaning – is the basis for immersion courses and other content-based instruction in second and foreign language programs (Bernhardt 1992; Genesee 1987; Lambert & Tucker 1972; Met 1994; Swain 1991). It is also the implicit justification for placing minority language children in classes where they are taught subject matter in their L2 – either in “submersion” (where most other students in the class are native speakers of the majority language) or in “sheltered” courses (where most students are L2 speakers and teachers make accommodation to their developing L2 proficiency).

Some SLA researchers and teachers have observed that while meaning-based courses effectively motivate learners and facilitate language acquisition, learners who do not also receive form-focused instruction fail to develop accuracy on certain linguistic features (Lightbown & Spada 1999; Lyster 1994; Swain 1988; L. White 1991). Experimental classroom studies have shown that the inclusion of form-focused instruction in meaning-based courses can contribute to interlanguage development and/or increased accuracy in learners’ use of the L2 (for review, see Norris & Ortega 2000). In these studies, both the presentation of rules or patterns and feedback on error have varied in terms of how explicitly learners’ attention was drawn

to the target forms. For example, J. White (1998) and Trahey and L. White (1993) provided enhanced input without explicit instruction or corrective feedback. Harley (1989) devised activities that required frequent use of the target forms in meaning-based activities but did not include explicit feedback on error or rule learning. Doughty and Varela's (1998) study emphasized "corrective recasts" in which the teacher drew students' attention, in both oral and written activities, to errors they made on the language patterns in focus. In the study by L. White, Spada, Lightbown, & Ranta (1991) learners received some explicit instruction, including metalinguistic guidance, and structured opportunities to practice and receive feedback on the targeted form. In reviews of research in both classroom and laboratory settings, Ellis (1995) and Spada (1997) concluded that the impact of instructional intervention is greater when it is more explicit in focusing learners' attention (see also Doughty 2001; Nicholas, Lightbown & Spada 2001).

The need for – and the effectiveness of – focused instruction may be determined in part by the extent to which errors impede communication. For example, errors in adverb placement (L. White 1991) or word order in questions (Spada & Lightbown 1993, 1999) do not normally cause communication breakdown. However misused possessive determiners (using his where her is intended, for example) have the potential of interrupting the flow of communication (J. White 1998). It may be that, in communicative interaction, errors that lead to communication failure will lead to negotiation of meaning and thus may not require focused instruction and corrective feedback. This has been suggested by several SLA theorists (Doughty & Williams 1998; Harley 1993), but to our knowledge, this hypothesis has not been the basis of systematic research in classroom settings. In any case, in a classroom where learners share the same L1, there is less likelihood of communication breakdown even on these features as students usually understand each other's interlanguage (Lightbown & Spada 2000). In such contexts, explicit instruction may be required to draw learners' attention to differences between their shared interlanguage and the L2. Some language features on which such intervention has been observed to be successful in ESL classes for francophone students include the distinction between be and have (Lightbown 1991) and between his and her (J. White 1998; J. White & Ranta 2002) and the placement of adverbs in simple sentences (L. White 1991).

Findings from classroom research with young learners, as well as SLA research and theory involving older learners, confirm that the features learners are most likely to have long term difficulty acquiring through communicative interaction are those in which there is a misleading similarity between the L1 and the L2 (Han & Selinker 1999; J. White 1998; Zobl

1980, 1985). This suggests that these are the L2 features that are most likely to require form-focused instruction. It has been hypothesized that, in order for learners to overcome their difficulties with these features, it may be necessary to provide instruction that is not only explicit with regard to the L2, but that also draws attention to the specific differences between L1 and L2 (Spada & Lightbown 1999).

In the published experimental research on classroom SLA, the instructional emphasis has usually been on the target language itself, rather than on explicit comparisons of the L1 and L2. Most of the instructional research which includes a contrastive L1/L2 component has been done with adult learners receiving traditional discrete-point grammar instruction (Sheen 1996; Von Elek & Oskarsson 1973). One study that included a contrastive component in communicative classrooms with adolescent school-aged L2 learners was carried out by Kupferberg and Olshtain (1996). They compared the learning outcomes of Israeli high school learners who were taught certain features of English using different varieties and combinations of communicative activities that also contained explicit contrastive information about Hebrew and English. The group who received the contrastive information performed better than the group who did not. J. White and Ranta (2002) found that the provision of contrastive information improved young francophone learners' performance on possessive determiners, but their study did not include a comparison group receiving explicit instruction without a contrastive component. We are not aware of other published studies of the extent to which young school-age L2 learners receiving communicative or content-based instruction can benefit from explicit contrasts between the L1 and L2 or of the benefits such contrasts may offer beyond those afforded by explicit instruction that is focused on the target language itself.

### Research questions and hypotheses

The question that motivated the present study was: Will explicit instruction that includes *contrastive* information about the L1 and L2 be more effective than explicit instruction without a contrastive component?

To investigate this question, we targeted two linguistic features that our previous research has shown to present persistent difficulties in the developing interlanguage of young francophone learners of L2 English. These features are the third person singular possessive determiners, his and her, and question formation, specifically subject/auxiliary inversion when the

subject is a noun or full noun phrase. We hypothesized that explicit form-focused instruction which includes a contrastive L1/L2 component would be more effective than instruction without a contrastive L1/L2 component. It will be seen that strict control of learners' exposure to instruction on the target features, including contrastive information, was not possible.

## **Linguistic features**

### *Possessive determiners*

Deciding between the possessive determiners his and her causes problems for many francophone learners of English. These problems appear to be due to differences in the way gender is assigned in the two languages. In French, the gender of masculine and feminine third person possessive determiners agrees with the grammatical gender of the possessed entity, while in English, it agrees with the natural gender of the possessor. That is, the French equivalent of English his may be either *son* (masculine) or *sa* (feminine), depending on the grammatical gender of the object possessed. Similarly, either *son* or *sa* is the equivalent of her (see J. White 1998, for further discussion).<sup>2</sup>

We can get a view of the problem by looking at contexts where the natural gender of the possessor is different from the natural gender of the entity possessed. These are referred to as *kin-different* contexts. In these contexts, learners who produce ungrammatical sentences like *He cry in the arm of her mother* and *It's a mother with his little boy* seem to be using the French rule. On the other hand, learners who produce a substantial number of sentences like *He cry when he see his mother* and *The mother help her son to get dress* appear to be applying the English rule for both his and her.

English and French also differ in the way they make reference to possession of body parts. In French, the possession of body parts may be referred to using the definite article, and possession is marked by a reflexive pronoun that agrees with the subject (*Alice se lave les cheveux*).<sup>3</sup> In English, possession of body parts is normally indicated with the same possessive determiners that refer to the possession of other objects or people (*Alice is washing her hair*).

Previous research with French L1 learners of English has documented a developmental pattern in the use of his and her in oral production among a large number of French-speaking L2 learners in classroom settings (Lightbown & Spada 1990; Martens 1988; Zobl 1985). J. White (1996, 1998) fine-

*Table 1.* Developmental sequence and stage assignment criteria in the acquisition of the English agreement rule for his/her by French-speaking learners

<b>Pre-emergence</b>	
Stage 1	<b>Avoidance of <u>his</u> and <u>her</u> and/or use of definite article</b> The little boy play with <u>the</u> bicycle. He have band-aid on <u>the</u> arm, <u>the</u> leg, <u>the</u> stomach.
Stage 2	<b>Use of <u>your</u> for all persons, genders and numbers, (minimum of 2 times; below criterion of 4 correct uses of <u>his</u> or <u>her</u>)</b> This boy cry in the arm of <u>your</u> mother. There's one girl talk with <u>your</u> dad.
<b>Emergence</b>	
Stage 3	<b>Emergence of either or both <u>his/her</u>, neither to criterion (4 correct uses)</b> A little boy do a cycle ride and he fall. He have a pain on back and butt. He said the situation at <u>her</u> mom.
Stage 4	<b>Preference for <u>his</u> or <u>her</u> (minimum of 4 instances); accompanied by over-generalization to contexts for the other form</b> The mother is dressing <u>her</u> little boy, and she put <u>her</u> clothes, <u>her</u> pant, <u>her</u> coat, and then she finish. The girl making hisself beautiful. She put the make-up on <u>his</u> hand, on <u>his</u> head, and <u>his</u> father is surprise.
<b>Post-emergence</b>	
Stage 5	<b>Differentiated use of <u>his</u> and <u>her</u> to criterion (4 correct uses), but not with kin- different gender</b> The girl fell on <u>her</u> bicycle. She look <u>his</u> father and cry. The dad put <u>her</u> little girl on <u>his</u> shoulder, and after, on <u>his</u> back.
Stage 6	<b>Differentiated use of <u>his</u> and <u>her</u> to criterion (4 correct uses); agreement rule applied to kin-different gender to criterion (2 correct uses) for either <u>his</u> or <u>her</u></b> The mother dress <u>her</u> boy. She put <u>his</u> pants and <u>his</u> sweater. He's all dressed and he say at <u>her</u> mother he go to the bathroom.
Stage 7	<b>Differentiated use of <u>his</u> and <u>her</u> to criterion (4 correct uses); agreement rule applied to kin-different gender to criterion (2 correct uses) for both <u>his</u> and <u>her</u>. The agreement rule may not be applied correctly to body parts.</b> The little girl fell the floor, and after she go see <u>her</u> father, and he pick up <u>his</u> girl in <u>the</u> arms.
Stage 8	<b>Error-free application of agreement rule to <u>his</u> and <u>her</u> to criterion (4 correct uses) in all contexts, including body parts</b> The little girl with <u>her</u> dad play together. And the dad take <u>his</u> girl on <u>his</u> shoulder and he hurt <u>his</u> back.

tuned earlier descriptions of the interlanguage development of this feature into a framework consisting of eight stages. She found that, in order to interpret the patterns in the data, it was useful to group the eight stages into three broader categories: Pre-emergence, Emergence, and Post-emergence.

In the Pre-Emergence stages, his and her do not appear in the learners' oral production. Possession is either not marked at all or is marked with an overgeneralized form, often your. The Emergence stages are characterized by a few instances of the target forms, or by the overgeneralization of either his or her. In the Post-emergence stages, learners gradually differentiate between his and her, using these forms correctly in a wider variety of linguistic contexts as they become increasingly accurate in their use of the English possessive determiner rule. Assignment to stages is based on an emergence criterion: a learner is assigned to the highest stage for which a small number of instances characteristic of that stage are produced. Thus, learners continue to make errors in contexts for PDs at all stages below Stage 8. Examples of learners' oral production at different stages and the criteria for assigning learners to stages are provided in Table 1.

In this study, we are interested in whether L2 learners who receive explicit contrastive information about possessive determiners (his and her) in English and who are shown how they differ from the French equivalents (*son* and *sa*) make more progress in their knowledge and use of possessive determiners than learners who receive explicit instruction that does not include a contrastive component.

### *Questions*

The acquisition of questions in English has also been observed to follow a developmental sequence (Pienemann, Johnston, & Brindley, 1988; Spada & Lightbown, 1993; Mackey & Philp, 1998). A summary of this sequence and examples from francophone learners' oral production are shown in Table 2.

In interpreting these stages, it is important to note that grammatical questions (e.g., What's that?) may occur in stage 1, but they are formulaic. Questions without inversion (stage 2) may be pragmatically appropriate in some contexts. Questions at stage 3 can be grammatical (e.g., Do you have a sister?) but they are assumed to be based on "fronting" without an understanding of inversion because learners also produce clearly ungrammatical questions with the same pattern of fronting (Do you can see a boy?). It is only when learners reach stage 4 and produce both wh-(with copula "be") and yes/no questions with a variety of auxiliaries that they are deemed to



Table 2. Developmental stages for question formation (adapted from Pienemann, Johnston, and Brindley 1988)

Stage 1	<b>Single words, sentence fragments, formulae</b>	Four children? A dog? What's that?
Stage 2	<b>Declarative word order – no inversion, no fronting</b>	It's a monster in the right corner? The boys throw the shoes?
Stage 3	<b>Fronting</b> <i>Wh</i> -fronting, no inversion  <i>Do</i> -fronting  <i>Other</i> -fronting	Where the little children are standing? What the dog are playing? Do you have a shoes on your picture? Does in this picture there is four astronauts? Is the picture has two planets on top?
Stage 4	<b>Inversion in <i>wh</i>- questions with copula</b>  <b>Inversion in yes/no questions (auxiliary other than <i>do</i>)</b>	Where is the sun?  Is there a fish in the water?
Stage 5	<b>Inversion in <i>wh</i>-questions with auxiliary and nonfinite verb (<i>do</i> and other auxiliaries)</b>	How do you say <i>proche</i> ? What's the boy doing? How many astronauts do you have?
Stage 6	<b>Complex questions</b> <b>Question tag</b> <b>Negative question</b> <b>Embedded question (de-inversion)</b>	It's better, isn't it? Why can't you go? Can you tell me what the date is today?

have grasped the inversion pattern that is grammatical in English. French-speaking learners, like learners from other language backgrounds, have been observed to acquire questions according to this developmental sequence.

To some extent, francophones can draw on their knowledge of French as they acquire English questions. Both English and French make use of subject-verb inversion to form questions. Many English and French questions are, with regard to the syntax of question formation, word-for-word equivalents. This is generally true of questions in which the subject is a pronoun and where the finite verb is an auxiliary, copula or modal.

Can you play outside?	<i>Peux-tu jouer dehors?</i>
Are you happy?	<i>Es-tu content?</i>
Have you seen my bicycle?	<i>As-tu vu mon vélo?</i>



Questions in the two languages also differ in several ways. The first is in the extent to which inversion is considered obligatory. In more formal registers, inversion is essentially obligatory for both yes/no and information questions in both languages.

Are they Italian?            *Sont-ils italiens?*  
Why is he angry?           *Pourquoi est-il fâché?*

In informal speech however, French makes much greater use of questions without inversion. English does this to some extent with yes/no questions, especially in certain conversational contexts. For *wh* questions, questions without inversion are ungrammatical in standard English, even in informal registers.

They're Italian?            *Ils sont italiens?*  
\*Why he is angry?         *Pourquoi il est fâché?*

Another difference is that French allows lexical verbs to be inverted with subjects, while English requires the addition of the auxiliary “do” to carry the information about tense and number.

*Viens-tu souvent au parc?*  
Do you often come to the park?  
(Literally: Come-you often to the park?)

French also has a form (*est-ce que*) that functions in some sense like do in that it can be placed at the beginning of a sentence to change a declarative to an interrogative. However, unlike do, the French form is invariant and does not change according to person, number or tense.

*Est-ce que nous sortons ce soir?*  
Are we going out tonight?  
(Literally: Is-it-that we are going out tonight?)

Finally, a difference that we have explored in previous research is that French, but not English, generally limits inversion to questions in which the subject is a pronoun.

\**Est Jean chez lui?*        Is John at home?

French requires the insertion of a dummy pronoun (noun subject-verb-dummy pronoun) to make the question grammatical.

*Jean est-il chez lui?*       \*John is-he at home?

There are exceptions to this pattern. For example, in certain information questions, the question word is followed by inversion, even when the subject

is a noun. When the verb is a lexical verb, the corresponding English sentence is ungrammatical.

*Où est le bureau du directeur?* Where is the principal's office?  
*Où va le professeur?* \*Where goes the teacher?  
*Que fait la petite fille?* \*What does the little girl?

We have observed that even when French-speaking learners reach stages 4 and 5, they continue to avoid, or to judge as ungrammatical, questions in which the subject-auxiliary inversion occurs with a noun (as opposed to a pronoun) subject. We have hypothesized that this is due to the influence of a constraint brought over from their L1 (See Spada & Lightbown 1999; Lightbown & Spada 2000).

In this study, we are interested in whether L2 learners who receive explicit contrastive information about question formation in English and French are able to make more progress in their knowledge and use of questions than learners who receive explicit instruction about question formation that does not include a contrastive component. Of particular interest is learners' ability to form questions and to recognize as grammatical, questions with inversion of noun subjects.<sup>4</sup>

### *Participants*

Four intact grade six intensive ESL classes in four different schools in the Montreal area participated in the study (N=approximately 90). Students (age 11–12 years) were beginning the second half of a five-month intensive course in which they were usually engaged in classroom activities with a focus on interaction and communication rather than on language itself. They had spent the first five months of grade six studying their academic subjects intensively in their mother tongue, French. The ESL component of the intensive year began in February and ended in June (for more information concerning intensive ESL in Quebec, see Lightbown & Spada 1994, 1997). This study was carried out in April and May, that is, the third and fourth months of their five-month program. All learners had studied English prior to grade six: learners in three of the classes had begun in grade 4, receiving between 90 and 120 minutes a week of meaning-focused instruction. Those in the fourth class had started English instruction in grade 5, receiving 90 minutes per week of the same meaning-focused curriculum and pedagogical approach.

*Research Design*

Learners in two classes received instruction on the possessive determiners his and her, and those in the other two classes received instruction on question formation, with an emphasis on inversion with noun subjects. The design of the study permitted each treatment group to serve as a control to the other. That is, the question (Q) group served as an uninstructed comparison group for the possessive determiner (PD) group and vice versa. The distribution of the participants is shown in Table 3.

In one of the PD classes (PD/CA), and one of the Q classes (Q/CA), the instructional intervention included a contrastive component to direct the learners' attention to the difference between their L1, French, and their L2, English. In the other class for each linguistic feature (PD/NCA and Q/NCA), explicit instruction was provided on the target feature, but no explicit reference was made to the L1/L2 contrast.

*Table 3.* Distribution of instructional treatments and target features

Class	Number in class*	Linguistic feature	+/- CA	First year of ESL
PD/CA	24	Possessive determiners	CA	5th grade
PD/NCA	27	Possessive determiners	No CA	4th grade
Q/CA	27	Questions	CA	4th grade
Q/NCA	18	Questions	No CA	4th grade

\*For some analyses, the numbers are smaller because learners were absent for a particular task or only a sample of learners participated.

All four teachers were highly proficient speakers of English although only one was a native speaker. All were ESL specialists and had taught intensive classes before. These four classes were selected for participation in the study following classroom observations. A short period of instruction was described, using the Communicative Orientation of Language Teaching (COLT) observation scheme (Spada and Fröhlich 1995). A general description of classroom interaction was based on COLT Part A; an adapted version of Part B focused on corrective feedback. This observation showed that, overall, the teachers devoted most of the classroom time to meaning-focused activity. However, we later learned that the Q/CA and PD/NCA teachers had drawn their students' attention to his and her at least briefly although

we are certain that they had not taught the target forms in any systematic way before we administered the pretests. The PD/NCA teacher told us that her students had considerable difficulty with pronouns, including possessives, and that this was the reason she had expressed interest in participating in the study. She also said that she gave her class “grammar homework” every night and spent about an hour the following day correcting it with the entire class. The content of the “grammar homework” was not focused on PD or Q patterns, although they may have appeared occasionally. We also discovered, after the study was under way, that the Q/CA teacher had drawn her students’ attention to question formation prior to the study, but we do not have any details about that instruction.

During the tenth week of the twenty-week intensive course, all four classes were pretested on both possessive determiners and questions. All learners completed paper and pencil tasks, and a stratified sample of learners also participated in oral tasks. This sample was selected with the assistance of the teachers, who were asked to identify learners from the strongest, weakest, and middle levels of proficiency in English among their students. Three or four participants were then randomly selected from each of these proficiency levels.

#### *Pedagogical treatment*

The experimental instructional activities were introduced in the week following the pretests. For four weeks, three times a week, learners in the PD and Q groups engaged in thirty-minute lessons consisting of a variety of whole-class games and activities that required them to understand and to produce the forms that were targeted in their treatment condition. The experimental lessons were developed by members of the research team, themselves experienced teachers or researchers in the context of intensive ESL classes. All of the lessons were taught by the regular ESL teacher in each class, and each teacher was given a complete package of materials, including sufficient copies of activity sheets for each learner and posters for the classrooms. During the first lesson, learners were taught a “rule of thumb”. For possessive determiners, the rule was *Ask “Whose is it? If it belongs to a man or boy, use his. If it belongs to a woman or girl, use her.”* For questions, the rule of thumb taught in the first lesson was *To form a yes/no question using can, will, is and are, invert the subject and the auxiliary verb.*

The metalinguistic component of the instruction was provided in two phases. During the first phase (the first week of the pedagogical interven-

tion), the Q classes practised inversion with pronouns but not nouns, and the PD classes practised examples that did not involve body parts. In the second phase (the next three weeks of the pedagogical intervention), Q groups practised inversion with nouns, as well as pronouns, and PD group activities included practice with body parts. For example, one day, Q groups played “Find someone who”, which required them to ask their classmates questions such as *Can you play hockey?* and *Are you taking piano lessons?* Another day, they played a game in which they answered questions like *Do beavers live in Canada?* and *Does your teacher speak Spanish?* Meanwhile, the PD groups played a game in which they had to use information provided by the teacher to talk about individuals in a large family so that their classmates could guess who was being described: *His mother was born in Italy;* and *Her father likes movies.* Another day, PD groups played a guessing game in which they had to describe their classmates without using their names: *His hair is short and his t-shirt is yellow* and *Her eyes are brown, her father is a teacher, and her brother is six years old.*

For all classes, the relevant rule of thumb (questions or possessive determiners), including some example sentences, was written on wall posters that were prominently displayed in the classroom. The posters were typographically enhanced in a way that was intended to increase the salience of the target features. For example, arrows were drawn to show the relationships of the linguistic elements that learners were expected to focus on. For possessive determiners, the arrows went from determiners to referents; for questions, they highlighted subject-verb inversion. In the classes receiving contrastive instruction, information about how the target feature is similar in French and English and how it is different was also included on the wall posters. These posters included example sentences in both languages, with similar typographical enhancement.

Throughout the instructional period, teachers provided corrective feedback when learners made errors on the forms targeted in the experimental activities and reminded them of the metalinguistic information, either by eliciting the rule of thumb or by pointing to the wall posters. It is important to emphasize that, outside the special activities, which occupied six hours of class time spread over four weeks, learners were involved in their regular, meaning-focused ESL activities.

Posttests were administered during the fifteenth week of the ESL course, immediately following the end of the pedagogical intervention. The research schedule is illustrated in Figure 1.

INTENSIVE ESL COURSE FEBRUARY – JUNE		
FEB – MARCH Weeks 1–10	APRIL – MAY Weeks 11–15	MAY – JUNE Weeks 16–20
Communicative ESL activities for approximately 4 hours per day	Experimental treatment period: 2 classes taught possessive determiners; 2 classes taught questions  Experimental instructional materials used for 30 minutes per day, 3 times per week; communicative activities continued for remainder of each school day (approximately 3.5 hours).	Communicative ESL activities for approximately 4 hours per day
Pretests in week 10	Posttests in week 15	General proficiency test (MEQ) in week 17

*Figure 1.* Research schedule

### Instruments

In the 17<sup>th</sup> week of the learners' intensive ESL program, a general proficiency test designed by the Ministère de l'Éducation du Québec (MEQ) was administered. Since 1986, we have used the MEQ test as a measure of general listening and reading comprehension for learners in intensive ESL. Its use is primarily to ascertain that the groups we choose for participation in experimental studies are typical of Quebec's primary intensive ESL students in terms of their overall knowledge of English.<sup>5</sup>

The pre/posttest measures of the linguistic features under investigation included paper and pencil metalinguistic tasks as well as written and oral production tasks for learners in both the Q and PD groups. All students completed all metalinguistic and written production tasks, and a sample of 11-12 learners per class participated in the oral production tasks. Those who participated in the oral production task also took part in a "meta-talk" interview regarding PDs.

*Possessive determiners: Passage correction*

In this metalinguistic task, learners read an illustrated story about a young boy's birthday party that contained 20 PD and 19 distracter errors. Seven of the PD errors were in *kin-different* contexts. Although they did not know how many errors the story contained, nor what kind they were, students were told that there was a maximum of one error per sentence and that there were no spelling errors. Learners were asked to read the story carefully, to put an X on any incorrect word they found, and to write the correct word above it. To ensure that they had understood the instructions, the researcher did three examples with the whole class on the blackboard. Similar tasks have been used with young learners in previous research on possessive determiners (Martens 1988; J. White 1996, 1998; J. White & Ranta, 2002).

*Possessive determiners: Oral production*

A version of this task has also been used in our previous research (J. White 1996, 1998; J. White & Ranta, 2002). Learners were asked to describe a series of six cartoon pictures representing family situations which offered contexts for the use of his and her, many of them kin-different (e.g. his mother, her brother). The task was administered to each learner individually, one picture at a time, and the interviews were audio-recorded and transcribed for analysis. Four of the pictures were the same at the pretest and posttest. To add variety and interest to the task, two of the pretest pictures were replaced with two different pictures at the posttest. Learners were prompted with the following: *We're going to look at some cartoon pictures of children and their parents. Here is the first one; can you tell me about this little boy/girl? If necessary, an additional prompt was used: What's the problem?* Since the interviewer's goal was to establish as many PD contexts as possible, especially those involving kin-different relationships, additional questions were often asked if the learner did not talk about the other people in the picture: *Who do you think this is?*

*Possessive determiners: Meta-talk interview*

Immediately following the oral production task, the researcher opened up the learner's passage correction task booklet to a pre-determined page and asked about some of the items that were corrected, as in the following extract from the data:

Here, Mathieu, I would like you to explain to me why you crossed out some words and changed them, o.k.? Here, for example, you crossed out this word and changed it.

The researcher also asked about items that were not corrected, both errors and correct forms:

Here, I know that you didn't cross anything out, but I would like you to think about it and tell me. "His father helped Susan to choose it." Would you change this sentence? Is there anything wrong with this sentence?

#### *Questions: Grammaticality Judgement*

This task is an adapted version of the explanation task reported in Lightbown & Spada (2000). Learners were presented with a list of 26 written questions and asked to judge whether they were correct. They were also asked to explain (in English or French) what was wrong with the questions they judged to be incorrect.<sup>6</sup>

#### *Questions: Written Production*

For the pretest, students were asked to imagine that their class would be going to an English camp during the summer vacation. They were to write 10 questions that they would like to ask the students from last year's English class who went to the same camp (e.g. How many people are in the cabin?). For the posttest, they were asked to write down 5 questions that a reporter might ask each of three interlocutors: (1) a resident of New York City; (2) a student in a high school they might someday attend, (3) the mother of Jacques Villeneuve, the race car driver.

#### *Questions: Oral Production*

In this task, a researcher seated across from a learner selected one card from each set of 4. A duplicate set was displayed in front of the student. For each picture set, the learner was invited to ask questions and try to guess which one the researcher was holding. Learners were encouraged to ask at least five questions, and if they hesitated, prompts were given (Spada & Lightbown 1993, 1999). Three different picture sets were used on the pretest. These were used again, and an additional picture set was included at the posttest. All interactions were audiorecorded, and the recordings were transcribed for subsequent analysis.



## Results

### MEQ Test

Results of the MEQ test were examined to determine whether the overall proficiency level – as measured by this test of listening and reading comprehension – was comparable for the four classes. An analysis of variance revealed that there was a significant difference among the groups  $F(3, 88) = 12.12, p < .001$ . The post hoc Tukey test showed that PD/CA performed less well than the other classes. There were no significant differences among the other three (See Table 4). PD/CA is the class that started learning ESL in grade 5, rather than grade 4. Although they proved to be the weakest class among those in this study, their mean score of 57% was within the range observed in other grade 6 intensive ESL classes.<sup>7</sup>

Table 4. MEQ test of reading/listening comprehension

Class	Mean score	Standard deviation	Standard error
PD/CA	57.48	13.35	2.78
PD/NCA	72.69	12.96	2.49
Q/CA	77.63	9.79	1.95
Q/NCA	72.61	12.28	2.98

### Possessive determiners: Passage correction

The passage correction task was analyzed with respect to each learner's accuracy in correcting PD errors at the pretest and again at the posttest. This was done by dividing the number of correctly corrected PD errors by the total number (20) of PD errors on the task.<sup>8</sup> Learners with an accuracy percentage below 50% were considered to be performing at a *low* level on this task. Learners who were between 50% and 75% accurate were considered to be performing at a *mid* level of accuracy. Learners whose overall accuracy was above 75%, and whose accuracy in correcting errors in *kin-different* contexts was above 50%, were considered to be at a *high* level of accuracy on this task.

The findings reported here are for the subset of learners who also completed the picture description task. Table 5 shows that learners in Q/CA outperformed those in the other three classes at the pretest: only two learners

out of eleven were below 50% accuracy in correcting PD errors, compared to nearly all the learners in the other classes.

Q/CA maintained this initial advantage over Q/NCA at the posttest. However, neither of the Q classes showed development over time, and four learners in Q/CA actually went *down*, three from high- to mid-levels of accuracy, and one from mid- to low-levels. In Q/NCA, nine learners showed no change, one went up from the low- to mid-level, and one went down, from the mid- to low-level. This is in marked contrast to the PD classes, both of which exhibited solid gains in accuracy after instruction on PDs. In PD/CA, six of the twelve learners who were at the low level at the pretest went up to mid level on the posttest. In PD/NCA, eight learners went up one level, three went up two levels, and only one stayed at the same level (low). Because of the small group sizes, the learners in PD/CA and PD/NCA were combined ( $n=24$ ), as were the learners in Q/CA and Q/NCA ( $n=22$ ), for the purpose of carrying out the Wilcoxon matched-pairs signed-ranks test. The pre- to posttest differences were not significant for the combined Q classes. However, for the combined PD classes, the pre- to posttest gains were significant ( $p < .001$ ).

Table 5. Passage correction: Percent of learners at different levels of accuracy in correcting possessive determiner errors

%	Q/CA (n=11)		Q/NCA (n=11)		PD/CA (n=12)		PD/NCA (n=12)	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Low-level	18	27	82	82	100	50	92	8
Mid-level	55	73	18	18	0	50	8	58
High-level	27	0	0	0	0	0	0	33

#### *Possessive determiners: Oral production*

Learners were assigned to oral development stages based on emergence rather than accuracy criteria. Following procedures that were developed in earlier PD studies (and explained above), the stages were grouped to reveal patterns in the learners' developing ability to use *his* and *her* and to differentiate between these two forms in varied linguistic contexts. Differentiation in kin-different contexts is required for assignment to the highest post-emergence stages (Stages 6, 7 and 8).

Table 6 shows further evidence that learners in Q/CA were more advanced with respect to PDs than learners in the other groups at the time of the pretest.

All of the learners in Q/CA were at an emergence or post-emergence stage, whereas a substantial proportion of the learners in the three other classes were at a pre-emergence stage at the time of the pretest. There is evidence of development from pre- to posttest in all classes. The performance of PD/CA is particularly noteworthy when we consider that the learners from this class, which had begun ESL instruction a year later than the others, was weaker than the other classes on the PD pretest measures as well as on the MEQ test. When the two Q classes and the two PD classes are combined, the Wilcoxon matched-pairs signed-ranks test shows significant improvement for both the Q group ( $p < .05$ ) and for the PD group ( $p < .0001$ ). The posttest distribution reveals that learners in the PD classes are clearly stronger than those in Q/NCA and comparable to those Q/CA, even though the PD classes performed less well than either of the Q classes on the pretest.

*Table 6.* Oral production: Percent of learners at each stage of PD development

%	Q/CA (n=11)		Q/NCA (n=11)		PD/CA (n=12)		PD/NCA (n=12)	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Pre-emergence (Stages 1-2)	0	9	37	9	75	17	50	8
Emergence (Stages 3-4)	73	27	45	73	25	17	33	25
Post-emergence (Stages 5-8)	27	63	18	18	0	67	17	67

Learners' spontaneous self-corrections on his and her as they completed the oral production task provide insight into their awareness of possessive determiners. While some learners in all groups self-corrected on possessive determiners at the pretest, learners in PD/NCA and Q/CA did so to a greater extent than learners in the other two classes, probably because their teachers had already drawn their attention to his and her before we began this study.

At the posttest, a high percentage of learners in both classes that were explicitly instructed on PDs, as well as learners in Q/CA, self-corrected in contexts for his and her, and they did this, on average, more than three times each during the task. Fewer learners in Q/NCA self-corrected. Furthermore, the number of self-correctors did not increase at the posttest,

and the self-correction rate for this class went down slightly. We interpret this as evidence that the Q/NCA teacher did not focus on his and her before or during the study.

Table 7. Oral production: Learners who self-corrected on his and her

	Pretest		Posttest	
	Number (and %) of learners	Mean number of self-corrections	Number (and %) of learners	Mean number of self-corrections
PD/CA (n=12)	2 (17%)	2	8 (67%)	3.5
PD/NCA (n=12)	8 (67%)	2	10 (83%)	3.3
Q/CA (n=11)	9 (82%)	2.4	8 (73%)	3.6
Q/NCA (n=11)	4 (36%)	1.75	4 (36%)	1.25

#### *Possessive determiners: Meta-talk interview*

A qualitative analysis of the transcribed meta-talk interviews was used in interpreting learners' performance on the other tasks. Of particular interest was the extent to which learners' awareness of a rule of thumb was reflected in their ability to judge grammaticality or to use possessive determiners correctly in their own oral production.

Example 1 comes from a learner in PD/NCA. In the interview, he used metalinguistic terminology, as well as the rule of thumb he had been taught, when asked about an error that he had not corrected on the passage correction task.

#### *Example 1.*

Researcher: Here you didn't make a change. "His father helped Susan to choose it."

Learner 22: Oops.

Researcher: Is that...do you still think that's correct?

Learner 22: No, I'm forget. I don't see it's her father.

Researcher: Why would you change it?

Learner 22: Because the possessor is um Susan and I say 'Who whose father is it?'

Researcher: Um hm.

Learner 22: It's Susan's father.

Researcher: So then you would say...

Learner 22: Her father.

Researcher: O.k., very good. Good explanation.

The same learner self-corrected six times on possessive determiners during the oral production task. In the long self-correction sequence in Example 2, he was describing a series of pictures in which a little boy falls off his tricycle and does not cry until he reaches his mother at home.<sup>9</sup>

*Example 2.*

Researcher: Can you tell me about this little boy and what's happening?

Learner 22: He do tricycle and he took a rock and he take uh and he fell uh after because he have a big...I don't know where is her his blessure (injury). Um he run and he stay her his bike here. He run inside his home and go ask at her at his Mom um uh talk about what he have and he fell down and he her mother her mother his mother comfort the boy.

Both the interview and the self-corrections suggest that this learner knows the English possessive determiner rule. However, he has not yet acquired control over the use of this knowledge under the constraints of real-time language processing. He overlooked many of the errors on the passage correction task, including those in kin-different contexts, and he met the kin-different criterion for stage 7 on the oral production task as a result of his self corrections.

Overall, more learners in the PD than in the Q group stated a rule of thumb or explicitly referred to possession and the correct gender distinction when asked to explain why they had (or had not) corrected certain of the possessive determiners in the passage correction task booklet. Learners in the Q classes were more likely to offer no explanation at all, or to demonstrate confusion or state the French rule. The learner seen in Example 3, from Q/NCA, got both the rule and the form backwards.

*Example 3.*

Researcher: And here it says "His father helped Susan to choose it", but you changed it to her father. Why did you put her? What's the reason?

Learner 5: Well, because the the uh...uh....o.k. because her is for the boy, no?

Researcher: Her is what?

Learner 5: For the boy. Masculine.

Researcher: O.k. Which boy?

Learner 5: The father is masculine.

Researcher: Masculine, o.k. So her is masculine.

The learner in Example 4, from Q/CA, also appears to be using the French rule in which the gender of the possessive determiner agrees with the grammatical gender of the entity possessed.

*Example 4.*

Researcher: O.k. What about this? “Susan is happy to be invited to the party and to wear his pretty new dress.” You changed his to her.

Learner 11: Because the dress it’s in the feminine and his is the masculine; it’s not a good match.

Example 5 comes from a learner in Q/NCA who was able to explain some of the errors he did and did not correct on the passage correction task although he remained at a low level of accuracy on this task at the posttest.

*Example 5.*

Researcher: So in this sentence “Susan is happy to be invited to the party and to wear his pretty new dress”. You changed it to “to wear her pretty new dress”. Why did you put her?

Learner 10: Because it’s a girl who put the dress, it’s Susan the dress. Is not she’s not a boy.

Researcher: O.k. And what about here in this sentence. “His father helped Susan to choose it”. Is that o.k. or would you change something?

Learner 10: Oh, uh, it’s a change here.

Researcher: You want to change this to what?

Learner 10: Her.

Researcher: Her father?

Learner 10: Yes.

Researcher: O.k. so you would make that change now if you could.

Learner 10: Yes.

This learner was at oral stage 7 at the pre- and posttests. At the pretest, he overgeneralized his and self-corrected three times; at the posttest, he overgeneralized her and self-corrected once. When we put these “snapshots” together, we see a learner who is actively working on the input data in an attempt to discover the English possessive agreement rule on his own. It may be that, had he received instruction on PDs at this point in his development, he would have been able to move forward more quickly.

*Possessive determiners: Summary*

In this study, there is evidence that students benefitted from instruction targeting PDs. On the passage correction task, PD groups showed solid gains in accuracy after instruction. Although students in the Q/CA group maintained their original advantage over Q/NCA, neither group showed significant development from pretest to posttest. On the oral task, both PD and Q groups

showed development over time with much stronger gains for the PD classes. The improvement of PD/CA is particularly striking because it was the weakest group at the pretest. Qualitative analyses of the meta talk interviews reveal that students in the PD groups are better able than those in the Q groups to explain how PD agreement works in English.

*Questions: Grammaticality Judgement*

Examination of learners' responses on this task indicated that although they had some tendency to accept more questions than they rejected, there was every indication that they were looking carefully at the sentences and making a judgement on each one. A repeated measures ANOVA showed that there was a significant improvement from pretest to posttest for all classes  $F(3,100) = 33.98, p < .01$ , but the difference between classes was not significant  $F(3,100) = 0.61$ .<sup>10</sup>

*Questions: Written production*

Each question that the learners produced on the pretest and posttest written production tasks was coded according to the developmental stages proposed by the Pienemann, Johnston and Brindley (1988) developmental framework (Table 2) which we have adapted for use with this population of learners (Spada & Lightbown 1993; Spada & Lightbown 1999). A second analysis was done to determine the frequency of inversion with noun and pronoun subjects.

Stage assignment was based on an emergence criterion. Students were assigned to the highest stage at which they produced two or more *different* questions. Table 8 compares the Q and PD classes in terms of the number of learners assigned to each stage of question development on the pretest and posttest. Overall, few or no learners were at stages 1, 2 or 5 at the pretest; most were at stage 3 or 4. No learners wrote stage 6 questions on either pretest or posttest. On both the pretest and the posttest, there were more stage 4 learners in the two Q classes than in the PD classes.

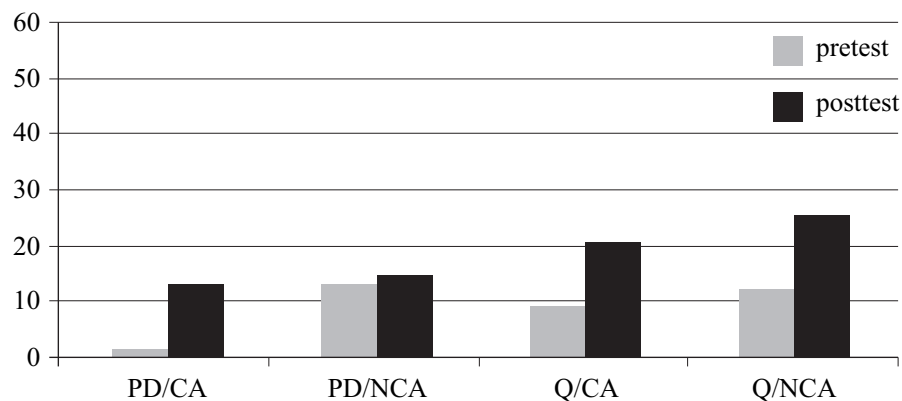
As shown in Table 8, 88% of the Q/CA learners were in Stage 4 at the posttest – more than in any other class. One cannot attribute the strong posttest performance of Q/CA to the experimental instruction, however, because they also had the strongest pretest performance. The other classes are a better test of the effect of instruction. A higher percentage of Q/NCA students were in stage 4 at the posttest than either of the PD classes. However, here again, it is difficult to attribute this to instruction alone as there were more students already in stage 4 at the pretest. All classes

improved, and four learners in PD/NCA moved to stage 5 although 40% of the students in this class remained in stage 3 at the posttest. Thus, on this measure, there is some evidence of a positive effect of instruction, outside the experimental intervention but learners who received instruction in question formation did not make substantially greater developmental progress than learners who received instruction in possessive determiners.

*Table 8.* Written production: Percentage of learners at each question stage, pretest and posttest

	Q/CA (n=25)		Q/NCA (n=17)		PD/CA (n=23)		PD/NCA (n=28)	
	%	Pre	Post	Pre	Post	Pre	Post	Pre
Stage 1	–	–	–	–	4	–	3	–
Stage 2	4	–	–	–	–	4	3	–
Stage 3	–	–	41	–	13	–	46	40
Stage 4	64	88	41	76	17	56	29	35
Stage 5	4	–	–	–	–	–	–	14
Unclassified	28	12	18	23	65	39	18	11

Note: “Unclassified” refers to learners who failed to produce at least two different non-formulaic utterances.<sup>11</sup>



*Figure 2.* Written Questions with Inverted Noun Subjects (Percent Inverted)

The second analysis of the written production data revealed some differences between the Q and PD groups that may be more plausibly attributed to the experimental instruction. The ten questions that each learner produced on the



pretest version of the written production task were examined. This resulted in a total of 430 questions for the Q classes and 510 questions for the PD classes. The 15 questions produced by each learner on the posttest version of the written production task were also examined: 630 questions for the Q classes and 765 questions for the PD classes. Figure 2 shows that relatively few learners wrote questions with subject noun inversion at either the pretest or the posttest. The highest percentage of questions with noun inversion at the pretest was 13% for both Q/NCA and PD/NCA. On the posttest, learners in all classes used more noun inversion, but the frequency in the Q classes and the magnitude of the change was greater than in the PD classes.

### *Questions: Oral Production*

For the sample of learners who participated in the oral question task, transcriptions were made of each pretest and posttest session. For each learner, the task resulted in the generation of approximately 15 questions for the three pictures used in the pretest and 20 questions for the four that were used at the posttest.

The first analysis assigned learners to a stage according to the highest stage at which they had at least two different, apparently non-formulaic questions (see Table 9).<sup>12</sup> There are two main findings from this analysis. First, 73% of the learners in the Q/CA class were already at stage 4 at the time of the pretest. This is the class with the highest pretest performance. Only one learner in this class progressed to stage 5, but by the time of the posttest, only one learner remained below stage 4. The strong pretest performance of this class makes it difficult to conclude that the strong posttest performance was due to the instruction. However, the other classes' results are more informative and do lend support to the hypothesis that focused instruction was beneficial in the development of questions.

In Q/NCA, 58% of the students were at stage 2 at the pretest. After the instruction on questions, 58% were at stage 4, and the performance of only two students was either below stage 3 or unclassified. Learners in PD/NCA and PD/CA were similar to learners in Q/NCA at the pretest, with over half in stage 2. On the posttest, only 4 learners in PD/CA and none in PD/NCA advanced to stage 4. The one learner in PD/NCA who had achieved stage 4 on the pretest dropped to stage 3 on the posttest and the three remaining learners did not have enough clearly non-formulaic questions for stage assignment and were categorized as "unclassified". These results from students' spontaneous oral performance provide some further evidence for the effectiveness of the instruction that was focused on questions.

*Table 9.* Oral production: Percentage of learners at each question stage, pretest and posttest.

%	Q/CA (n=11)		Q/NCA (n=12)		PD/CA (n=11)		PD/NCA (n=12)	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Stage 1	–	–	–	–	9	–	–	–
Stage 2	–	–	58	8	55	36	75	25
Stage 3	18	9	33	25	–	27	17	50
Stage 4	73	82	8	58	18	36	8	–
Stage 5	–	9	–	–	–	–	–	–
Unclassified	9	–	–	8	18	–	–	25

“Unclassified” refers to learners who produced fewer than two different non-formulaic questions.<sup>13</sup>

### *Questions: Summary*

The pedagogical treatment that was provided to young francophone students in intensive ESL classes led to some progress in terms of both knowledge and use of question forms. The difference between the students taught PDs and those taught Qs was not great, however, and on some measures, the PD students performed as well as the Q group at posttest. Furthermore, because of pretest differences, it was often not possible to conclude that between-class differences at the posttest were due to the instruction itself rather than to natural patterns of L2 development that would have occurred without intervention.

### **Discussion**

The primary question that we set out to investigate in this study was whether the provision of explicit *contrastive* information about the L1 and L2 is more effective than explicit instruction without a contrastive component in the development of language features known to be particularly problematic because of L1 influence. This question cannot be answered by our findings. The primary reason for this is the difficulty of controlling variables in research carried out in intact classrooms. The performance of learners in Q/CA was much stronger than that of the other classes on both linguistic features at the pretest. Thus, it would not be reasonable to conclude that their superiority on questions at the posttest was due to their having

received contrastive rather than non-contrastive instruction. As for the teaching and learning of PDs, we discovered that the teacher in the PD/NCA class had taught her students about PDs and given them the English rules that determine the choice of his or her prior to our experimental intervention. Although there was no evidence that she had systematically contrasted the L1 and L2 rules for PD use, there is no doubt that she had drawn the learners' attention to these differences. She specifically mentioned having told the students that, in using his for her (or vice versa) in kin-different contexts, they were making "an intelligent error", applying the rule that worked in French as they tried to acquire English. Thus we cannot conclude that the PD/CA had a CA component in their instruction while PD/NCA did not. Furthermore, the relatively advanced level already attained by Q/CA at the pretest may have been due to their teacher's prior attention to this feature although we have less information about this teacher's focus on it.

The results of this study do lend further support to the claim that explicit instruction – including that which is somewhat "obtrusive" (Doughty & Williams 1998) – can make a useful contribution within communicative language teaching. Q/NCA, which was far less advanced in questions than Q/CA at the pretest, improved more than the PD classes on some measures of question use. It is also evident, however, that instruction, as operationalized in this study, led to more substantial changes in learners' knowledge and use of possessive determiners than questions. There are several possible reasons for this difference: the importance of the form/meaning relationship represented by the two linguistic features, the relative complexity of "getting it right" for possessive determiners and questions, and the relative difficulty of assessing learners' progress in their knowledge and use of the features.

The observation that different target features respond to form-focused instruction in different ways has been reported in other quasi-experimental classroom research. For example, Lyster (1994) reports positive effects for form-focused instruction on the acquisition of the *tu/vous* distinction in French but not for the other more structurally and semantically complex features examined in his study. The finding that the his/her distinction in our study was more strongly influenced by instruction than were questions is consistent with the hypothesis that although instruction may be helpful in the acquisition of vocabulary, it is not effective for the acquisition of syntax (Schwartz 1993).<sup>14</sup>

Regardless of whether the feature is lexical or syntactic, learners may be able to benefit more from form-focused instruction "when the form in

question is crucial to the meaning being conveyed” (Lightbown 1998: 192). Errors in the word order of questions are essentially errors in *form* and do not seriously interfere with *meaning*. In contrast, the use of an incorrect possessive determiner can lead to genuine communication breakdown. In a sense, it may be said that the PD instruction focused on helping learners to *understand* what they heard and to *say what they meant* while the Q instruction focused on getting students to alter the form of questions that were already comprehensible. When learners understand that choosing the wrong form can lead to miscommunication, there is a motivation to make a change. The motivation to be accurate with regard to form may not be as strong. This may account in part for the progress seen in PDs among learners who did not receive instruction on this feature in the course of their communicative activities. That is, even in these classrooms where students shared the same interlanguage, there may have been instances where a misunderstanding led to negotiation that would have drawn their attention to PDs. Such misunderstandings would not be likely to occur when questions were asked without subject-verb inversion.

Other SLA theorists and researchers have predicted and/or found differences in the effectiveness of instruction on different linguistic features (e.g. Doughty & Williams 1998; Hulstijn & DeGraff 1994; VanPatten 1994). In laboratory studies, Robinson (1996) describes differential effects of explicit instruction on ‘simple’ versus ‘complex’ rules and DeKeyser (1995, 1996) and DeKeyser and Sokalski (2001) report interactions between instructional type and morphological complexity of target features. Although there are many unanswered questions, there does seem to be evidence that complex linguistic features are affected more by exposure to input containing the features than by explicit metalinguistic instruction. Questions certainly fall in the category of language features to which classroom learners have high-frequency exposure, and the progress made on question formation by students in this study who did not receive any explicit instruction on this feature may be attributable to the natural course of development as students continued to hear many correctly formed questions of all types in their teachers’ classroom language.

The change that is required to revise the incorrect use of a PD is essentially lexical, that is, the choice of the correct word for the intended meaning. To be sure, there is processing complexity inherent in PDs. Information present in the non-linguistic context must be monitored and/or the information in one part of the sentence (e.g., the subject) must be monitored when selecting the correct form in another part of the sentence (e.g., the verb complement or direct object). Nevertheless, once the learner has

begun to use both his and her in some contexts, the essential task is to choose between them and that choice does not require a change in the structure of the sentence. This contrasts with question formation, which requires changes to several sentence elements, including subject/auxiliary inversion and in some instances the insertion of auxiliary do to carry information about tense and number (see also Pienemann 1999, for discussion of processing complexity).

The difficulty of measuring the effect of instruction in this study arises from two factors. First, as we noted above, the learners in Q/CA were already quite advanced in their use of questions at the pretest. As a result, they had less room to grow. Most of them were already at stage 4, and the questions they asked were almost always correct and appropriate. It may be, however, that their “failure” to move to stage 5 or the “failure” of learners in Q/NCA to move out of stage 3 is more apparent than real. It may be due to the difficulty of determining with certainty that there is an obligatory context for a particular type of question. It will be recalled that even though learners in Q/CA received no systematic instruction on PDs, they progressed in their knowledge and use of PDs. Measuring PD progress was relatively straightforward. For example, if a learner says, “The little boy talk to *her* mother” in a context requiring “*his* mother”, the error is immediately apparent, and it is possible to locate the point on a developmental continuum that the student has reached. However, there is great variability among the possible forms that questions can take when learners have choices. In the oral and written production tasks for this study, students could ask “Do you have a boy beside a tree?” (stage 3), “What is beside the tree?” (stage 4) or “What do you see beside the tree?” (stage 5). All questions are essentially correct, and all are appropriate for moving the guessing game along. If a learner failed to produce any stage 5 questions, we could not know with certainty that this was because he or she was unable to do so. We may therefore have failed to see progress even when it had occurred.

## **Conclusion**

This study has provided additional evidence for the effectiveness of explicit form-focused instruction in helping students in CLT contexts increase their knowledge and use of language features that are difficult in part because of the misleading similarity between the French and English forms to express the same meanings. While some of the pretest/posttest changes appear to be attributable to natural developmental progress and to instruction provided

by the teachers outside the experimental intervention, it is possible to conclude that the experimental teaching contributed to the rate and the quality of students' increased ability to use these difficult linguistic features. The impact of instruction on students' ability to use PDs was more dramatic than the apparent effect of instruction on their knowledge and use of questions.

Long's "interaction hypothesis" (1996) and Swain's "output hypothesis" (2000) suggest that when learners engage in communicative interaction, they may encounter communication breakdown in which they are forced to notice that they are not making themselves clear. In such contexts, they may also notice that their language use differs from that of more proficient speakers or from that of other L2 speakers, as would be the case in a linguistically heterogeneous ESL class. However, many interlanguage features do not interfere with meaning, even though they diverge from the target version of the feature. Questions without inversion are usually completely comprehensible to native speakers of English or interlanguage speakers from different L1 backgrounds. Furthermore, we have argued that, in CLT situations where learners share the same L1, they may have similar interlanguage patterns, rendering their non-target like utterances comprehensible to their peers. This would be the case for incorrect gender selection for possessive determiners. The evidence suggests that, among these young francophone learners, interlanguage versions of both questions and possessive determiners will be hard to change without explicit instructional intervention. Both features have been shown, in this study and previous ones, to be responsive to instruction. What this study also suggests, however, is that even with explicit instruction, learners are more likely to make changes in their use of language features that have an impact on meaning.

One important aspect of this research was the use of a variety of assessment instruments and procedures in probing learners' knowledge of the features that had been taught. In classroom SLA research, it is important to assess learners' performance on measures in which attention is not focused on the target features but is more spontaneous in nature. In this study, both the oral and written production tasks encouraged students to focus on meaning. The findings from those measures add to the evidence from the metalinguistic tasks that explicit form-focused instruction, provided within the overall framework of communicative language learning, promotes progress in the acquisition of a second language.

## Notes

1. The research reported here was supported by a grant to Nina Spada and Patsy M. Lightbown from SSHRC (the Social Science and Humanities Research Council of Canada). We are grateful to the teachers and students who cooperated enthusiastically in this study. We are deeply indebted to Ahlem Ammar, Christine Brassard, Astrid Liden, and Lucy Fazio for their important contributions to the research, and to Randall Halter, who provided indispensable assistance from beginning to end.
2. When the possessed entity is plural (whether masculine or feminine), the equivalent of both *his* and *her* would be *ses*. In the present study, all examples in the instructional materials involved the use of singular nouns.
3. Reflexive pronouns agree with the subject in person and number. However, in the third person the same form (*se*) is used for singular and plural. There is no gender distinction in the pronoun.
4. Picard (2002) has argued that differences between learners' acquisition of inversion in questions with noun and pronoun subjects cannot be due to L1 influence because "with the exception of [some] restricted cases..., pronoun-verb inversion in question formation is not part of the active grammar of Canadian French" (p.66). It should be recalled, however, that these students have been taught the rules for inversion in questions in standard French in their French language arts classes. There is no reason to assume that they would draw exclusively on the informal grammar of their oral French in their developing knowledge of English. Their reliance on more formal varieties of French might be particularly influential when they are engaged in paper and pencil tasks. See Lépine (2001) for evidence of metalinguistic awareness of rules governing inversion in English and French in francophone secondary school students.
5. This test was originally designed more than twenty years ago. It was administered to many thousands of learners in the middle of secondary school in order to assess the knowledge of English that they might have acquired either in or out of their ESL classes. The test is somewhat dated in terms of both its content and its format, and it is not possible to say that it is a "secure" test. Nonetheless, it has the value of permitting us to compare each new group of learners against a baseline made up of approximately one hundred intensive ESL groups at the grade 5 and 6 levels.
6. The explanations are not included in this paper. Very few students wrote any explanation of their judgements.
7. It will be recalled that the MEQ test was administered in the 17<sup>th</sup> week of the intensive ESL course, after the experimental intervention had ended. This was done so that the results of the test would be comparable to those obtained in the many intensive ESL classes that had been tested previously. For this reason, we did not know how students performed on this test until the experiment was over.



8. The distracter items were analyzed only to determine that students were attending to errors and making changes in a way that confirmed that they understood the task. No quantitative analysis of the distracter errors was performed.
9. The final form produced in the self-correction sequence is the one that was considered in assigning the learner to a stage.
10. This analysis was based on the percentage of students who judged each question correctly on the grammaticality judgement task. Thus, the degrees of freedom for the repeated measures reflects four classes ( $df = 3$ ) and the 26 items on the task ( $df = (26 \times 4) - 4$ ) rather than the number of students in each class.
11. Assigning learners to a particular stage required that they produce at least two different questions appropriate to that stage which were not apparently formulaic. In the written production task, most of the learners who were categorized as “unclassified” were those who wrote only one question formula, repeatedly writing, for example, “Can I...”, “Can we...”, “Can you...” or “Is it...”. See further discussion of formulaic questions in notes 12 and 13.
12. Assigning learners to a particular stage required that they produce at least two different questions appropriate to that stage. For example, although “How do you say *trottoir*?” is technically a stage 5 question, it was treated as a “chunk” if it was the only stage 5 question a learner asked. Similarly, if all a learner’s stage 4 questions were of the type “Where is (the dog)?”, the learner was not deemed to have demonstrated the ability to create stage 4 questions but was rather assumed to have memorized a formula.
13. For most learners, stage assignment was straightforward, but a few produced no questions at all that were unequivocally non-formulaic. For example, some produced *only* “Do you have...” (stage 3) and/or “where is...” (stage 4) questions. These are the students called “unclassified” in Table 9.
14. On the other hand, N. Ellis (1994) has argued that different aspects of vocabulary are processed differently and may respond to different types of instruction in different ways. For example, surface forms of vocabulary (e.g. pronunciation elements and orthographic combinations) may be learned implicitly while the meaning and referential aspects of words may be learned explicitly.

## References

- Bernhardt, E. (ed.)  
 1992 *Life in language immersion classrooms*. Clevedon: Multilingual Matters Ltd.
- DeKeyser, R.  
 1995 Learning second language grammar rules: An experiment with a miniature linguistic system. *Studies in Second Language Acquisition* 19: 249–297.  
 1997 Beyond explicit rule learning: Automatizing second language morphosyntax. *Studies in Second Language Acquisition* 19: 195–221.



- DeKeyser, R. and K. Sokalski  
2001 The differential role of comprehension and production practice. *Language Learning* 51: Supplement 1. Form-focussed instruction and second language learning, 81–112.
- Doughty, C.  
2001 Cognitive underpinnings of focus on form. In *Cognition and second language instruction*, P. Robinson (ed.), 206–257.
- Doughty, C. and E. Varela  
1998 Communicative focus on form. In C. Doughty and J. Williams (eds.), *Focus on form in classroom SLA*, 114–138. Cambridge: Cambridge University Press.
- Doughty, C. and J. Williams  
1998 Pedagogical choices in focus on form. In *Focus on form in classroom SLA*, C. Doughty and J. Williams (eds.), 114–138. Cambridge: Cambridge University Press.
- Ellis, N. C.  
1994 *Implicit and explicit learning of languages*. London: Academic Press.
- Genesee, F.  
1987 *Learning through two languages: Studies of immersion and bilingual education*. Cambridge: Newbury House Publishers.
- Han, Z-H. and L. Selinker  
1999 Error resistance: Towards an empirical pedagogy. *Language Teaching Research* 3: 248–275.
- Harley, B.  
1989 Functional grammar in French immersion: A classroom experiment. *Applied Linguistics* 10: 331–359.
- Hulstijn, J. and R. DeGraff  
1994 Under what conditions does explicit knowledge of a second language facilitate the acquisition of implicit knowledge? A research proposal. *AILA Review* 11: 97–112.
- Kupferberg, I. and E. Olshtain  
1996 Explicit contrastive instruction facilitates the acquisition of difficult L2 forms. *Language Awareness* 5: 149–165.
- Lambert, W. and G. R. Tucker  
1972 *Bilingual education of children: The St. Lambert experiment*. Rowley, MA: Newbury House Publishers.
- Lépine, C.  
2001 *Adolescent learners' awareness of first language influence on their second language knowledge*. M.A. diss., McGill University, Montreal.
- Lightbown, P. M.  
1998 The importance of timing in focus on form. In *Focus on form in classroom second language acquisition*, C. Doughty and J. Williams (eds.), Cambridge: Cambridge University Press.

- Lightbown, P. M. and N. Spada  
 1990 Focus-on-form and corrective feedback in communicative language teaching: Effects on second language learning. *Studies in Second Language Acquisition* 12: 429–48.  
 1994 An innovative program for primary ESL in Quebec. *TESOL Quarterly* 28: 563–79.  
 1997 Learning English as a second language in a special school in Québec. *Canadian Modern Language Review* 53: 315–355.  
 1999 *How languages are learned*, revised edition. Oxford: Oxford University Press.  
 2000 Do they know what they're doing? L2 learners' awareness of L1 influence. *Language Awareness* 9: 198–217.
- Long, M. H.  
 1996 The role of the linguistic environment in second language acquisition. In *Handbook of Second Language Acquisition*, W. C. Ritchie and T. K. Bhatia (eds.), 413–468. San Diego, CA: Academic Press.
- Lyster, R.  
 1994 The effect of functional-analytic teaching on aspects of French immersion learners' sociolinguistic competence. *Applied Linguistics* 15: 263–287.
- Mackey, A. and J. Philp  
 1998 Conversational interaction and second language development: Recasts, responses and red herrings. *Modern Language Journal* 82: 338–356.
- Martens, M.  
 1988 Recognition and production of pronouns by Francophone learners of English as a second language. Unpublished master's thesis, Concordia University, Montreal.
- Met, M.  
 1994 Teaching content through a second language. In *Educating second language children: The whole child, the whole curriculum, the whole community*, F. Genesee, (ed.). Cambridge: Cambridge University Press.
- Nicholas, H., P. M. Lightbown and N. Spada  
 2001 Recasts as feedback to language learners. *Language Learning* 51: 719–758.
- Norris, J. and L. Ortega  
 2000 Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning* 50: 417–528.
- Picard, M.  
 2002 L1 interference in second language acquisition: The case of question formation in Canadian French. *IRAL* 40: 61–68.
- Pienemann, M.  
 1999 *Language processing and second language development*. Amsterdam: John Benjamins.

- Pienemann, M., M. Johnston and G. Brindley  
1988 Constructing an acquisition-based procedure for second language assessment. *Studies in Second Language Acquisition* 9: 92–122.
- Robinson, P.  
1996 Learning simple and complex second language rules under implicit, incidental, rule-search and instructed conditions. *Studies in Second Language Acquisition* 19: 233–247.
- Schwartz, B.  
1993 On explicit and negative data effecting and affecting competence and linguistic behavior. *Studies in Second Language Acquisition* 15: 147–162.
- Sheen, R.  
1996 The advantage of exploiting contrastive analysis in teaching and learning a foreign language. *IRAL* XXXIV: 183–197.
- Spada, N.  
1997 Form-focussed instruction and second language acquisition: A review of classroom and laboratory research. *Language Teaching Abstracts* 30: 73–87.
- Spada, N. and M. Fröhlich  
1995 *The Communicative Orientation of Language Teaching COLT observation scheme: Applications for research and teacher education*. Sydney: National Centre for English Language Teaching and Research NCELTR.
- Spada, N. and P. M. Lightbown  
1993 Instruction and the development of questions in L2 classrooms. *Studies in Second Language Acquisition* 15: 205–224.  
1999 Instruction, L1 influence and developmental readiness in second language acquisition. *Modern Language Journal* 83: 1–22.
- Swain, M.  
1988 Manipulating and complementing content teaching to maximize second language learning. *TESL Canada Journal* 6: 68–83.  
1991 French immersion and its offshoots: Getting two for one. In *Foreign language acquisition: Research and the classroom*, B. Freed (ed.), Lexington, MA: Heath.  
2000 The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In *Sociocultural Theory and Second Language Learning*, J. Lantolf (ed.), 97–114. Oxford: Oxford University Press.
- Trahey, M. and L. White  
1993 Positive evidence and preemption in the second language classroom. *Studies in Second Language Acquisition* 15: 181–204.
- VanPatten, B.  
1994 Examining the role of consciousness in second language acquisition: Terms, linguistic features and research methodology. *AILA Review* 11: 27–37.

- VanPatten, B.  
1996 *Input processing and grammar instruction: Theory and research*. Norwood, NJ: Ablex.
- Von Elek, T. and M. Oskarsson  
1973 *Teaching foreign language grammar to adults: A comparative study*. Stockholm: Almqvist and Wiksell.
- White, J.  
1996 *An input enhancement study with ESL children: Effects on the acquisition of possessive determiners*. PhD diss., McGill University, Montreal.  
1998 Getting learners' attention: A typographical input enhancement study. In *Focus on form in classroom second language acquisition*, C. Doughty and J. Williams (eds.). Cambridge: Cambridge University Press.
- White, J. and L. Ranta  
2002 What you know and what you do: The relationship between meta-linguistic performance and oral production in a second language. *Language Awareness* 11 (4): 259–290.
- White, L.  
1991 Adverb placement in second language acquisition: Some effects of positive and negative evidence in the classroom. *Second Language Research* 7: 133–161.
- White, L., N. Spada, P. M. Lightbown and L. Ranta  
1991 Input enhancement and L2 question formation. *Applied Linguistics* 12: 416–432.
- Zobl, H.  
1980 The formal and developmental selectivity of L1 influence in L2 acquisition. *Language Learning* 30: 43–57.  
1985 Grammars in search of input and intake. In *Input in second language acquisition*, S. Gass and C. Madden (eds.). Rowley, MA: Newbury House.

## Structure complexity and the efficacy of explicit grammar instruction

*Alex Housen, Michel Pierrard and Siska Van Daele*

Previous research indicates that the effects of explicit instruction for second language acquisition depend on the complexity of the target feature but the exact relationship between complexity and instruction remains unclear. The first part of this paper discusses the concepts of *explicit instruction*, *linguistic structure* vs. *pedagogical rule* and *complexity*. The second part describes an experiment which examines the effects of explicit instruction on second language learning as well as the impact of complexity on the effectiveness of explicit instruction. Complexity is defined in terms of the *functional markedness* of the linguistic structure as proposed by Givón (1995). Two grammatical target structures are thus selected: French sentence negation (less marked/simple structure) and French passive constructions (more marked/complex structure).

Subjects are 69 Dutch-speaking learners of French-Foreign Language (14–15 years old) in three intact classes. One class received explicit instruction on the simple structure, the second class received explicit instruction on the complex structure and third class received no instruction on either structure. Differences between the groups in their knowledge of the target structures were measured using a pretest-posttest design. Each test consisted of three tasks varying in the degree of planning and activation of explicit vs. implicit language knowledge: a grammaticality judgement task, a controlled written production task, and an unplanned oral production task.

A first series of analyses reflects a clear positive effect of explicit instruction on learners' mastery of the target structures. The strongest effect is found in the learners' unplanned oral production. This would suggest that explicit instruction promotes not only explicit grammatical knowledge as shown by previous studies, but also implicit knowledge.

The second series of analyses shows that learners who were taught the complex structure systematically gain more from the explicit instruction than do learners who were taught the simpler structure. However, this trend is merely suggestive rather than significant as the statistical analyses identified no unequivocal effects of the complexity of the target structures on the efficacy of the instruction provided.

## 1. Introduction

The study presented in this chapter is part of a larger research program at the Vrije Universiteit Brussel designed to investigate whether grammar instruction leads to higher levels of L2 competence and L2 proficiency and, if so, which factors influence the effectiveness of grammar instruction. These questions are pursued in a series of quasi-experimental studies with Foreign Language learners in the Belgian (secondary education) system (e.g. Hendrix, Housen & Pierrard 2002, 2003). Our intention is to provide learners in different classes with instruction in the use of specific grammatical forms chosen on the basis of linguistic theory and other SLA research, and to investigate whether the instructional treatments lead to differences in learners' performance in a variety of language tasks. The purpose of the study presented in this chapter is to investigate the role of *explicit* (or *formal*) grammar instruction, here defined as the provision of metalinguistic descriptions and explanations of grammatical features.<sup>1</sup>

Explicit grammar instruction is an important component of many instructed second language learning contexts but its role in SLA has been controversial since antiquity (Richards & Rogers 1986). Even today the question remains whether the explicit teaching of rules really enables or facilitates SLA. Language practitioners have (often tacitly) assumed that it does. The research community, however, is more divided, and the efficacy of explicit grammar teaching has been questioned both on theoretical and empirical grounds.

Some SLA theorists assume that explicit instruction can only lead to conscious, verbalizable metalinguistic knowledge, but not to the kind of implicit competence that underlies fluent, spontaneous language use (e.g. Krashen 1985, 1993). Other theorists, however, see a more constructive role for this type of instruction by allowing for an *interface* between explicit and implicit knowledge (e.g. Robinson 1995; Hulstijn 1989, 1995; Sharwood Smith 1991; Schmidt 1990, 1993, 1995; N. Ellis 2002). Empirical research on the effects of explicit instruction on learners' knowledge and use of grammatical structures has also yielded mixed results. Some studies found explicit explanations of linguistic structures to be beneficial to L2 learners' control over these structures (Spada & Lightbown 1993; Carroll & Swain 1993; Allen et al., 1990; Alanen 1995; Robinson 1995, 1996, 1997; Willams & Evans 1998). Other studies, however, found no such effects (e.g. White 1990; Felix & Weigl 1991; Bley-Vroman 1989). Ellis (1994) referred to this state of affairs as "the paradox of formal language instruction": "formal instruction results in faster and more successful language

learning and yet it often fails to teach learners specific linguistic features” (Ellis 1994: 107). Experimental research conducted over the last ten years suggests that the effectiveness of explicit instruction depends not on one but on a variety of factors, including the type of learner, the type of language knowledge or aspect of language proficiency considered, the type of instruction provided and the nature of the grammatical feature taught (Norris & Ortega 2000; Hulstijn & De Graaff 1994; Ellis 2001; DeKeyser 1998). Our own research deals with the last factor.

The impetus for our investigation is the belief that not every type of grammatical feature makes an equally good candidate for explicit instruction. Statements to this effect have repeatedly been made, but insufficiently demonstrated in the SLA literature, starting with Krashen (1982) who introduced the distinction between “rules that are easy to acquire [implicitly] but hard to learn [via explicit instruction]” and “rules which are easy to learn but hard to acquire”, the implication being that the latter type of rules are the primary candidates for explicit instruction. Although the implication may be clear, the problem remains to determine exactly what makes a rule hard to acquire but easy to learn and, hence, effective for instruction. Factors that have been proposed in this respect include the complexity, perceptual salience and communicative load of the underlying rule structure, the degree of contrast with the equivalent rule in the L1, the clarity, intelligibility and memorability of the formulation of the rule and the way the rule is married to examples (DeKeyser 1998; Harley 1994; Doughty & Williams 1998; Ellis 1997).

This chapter focuses on the *complexity* factor. Although several authors have argued that the role of explicit instruction for second language acquisition depends crucially on the *complexity* of the target feature, the exact link between complexity and instruction remains unclear. Basically there are two, seemingly contradictory views: those who argue that instruction should focus on simple rules (e.g. Pica 1985) and those who argue that instruction should focus on complex rules (e.g. De Graaff 1997). Part of the contradiction lies in the fact that different researchers have interpreted ‘instruction’, ‘complexity’ and ‘rule’ in different ways. Before presenting our own study, we will therefore first explain how we have defined and operationalized these terms. We also briefly situate our approach to some explicit proposals made in this regard by Hulstijn & De Graaff (1994), DeKeyser (1998) and Dietz (2002).

## 2. Conceptual and terminological issues

### 2.1. Explicit instruction

It is clear that the term (*second/foreign language*) *instruction* (or: *teaching*) covers a wide range of methods, approaches, pedagogical principles, strategies, techniques and didactic activities. For an extensive discussion of the term and for language instruction typologies, see Stern (1992), Norris & Ortega (2000), Long & Robinson (1998), Ellis (1997, 2001), among others. We refer to the type of instruction used in this study as *explicit* instruction. By explicit instruction we understand any pedagogical activity that meets the following criteria:

- There is an attempt to isolate a specific linguistic feature for focused attention;
- Learners are provided with data which illustrate the target feature;
- Learners are asked to perform one or more operations on the data;
- Learners are expected to make some intellectual effort to consciously understand the target feature;
- Learners are presented with a metalinguistic rule describing the formation and/or use of the target feature, or they may be required to verbalize such a rule.

Our notion of explicit instruction corresponds to what others have called *formal instruction* (Ellis 1994) or *Focus-on-FormS instruction* (Long 1991; Long & Robinson 1998; see also Sheen, this volume). Further characterization of explicit instruction will be provided later when we describe the instructional treatment used in our study (section 3.4.1.).

### 2.2. Linguistic structures versus pedagogical rules

In line with proposals of other researchers (e.g. Hulstijn & De Graaff 1994; De Graaff 1997; Dietz 2002) we distinguish between (a) ‘rule’ as the *linguistic feature* or *phenomenon* whose acquisition by the learner is the goal of the instruction, and (b) ‘rule’ as the *metalinguistic description* or *explanation* which is provided to the learner as part of the instructional activity intended to promote the acquisition of the target feature. We shall refer to the first meaning of ‘rule’ as *structure*. The term *linguistic structure* then is used to refer to the symbolic constructs postulated by linguists to denote or model



observable linguistic phenomena (e.g. patterns of structural co-variance and form-function mappings) and/or their underlying mental representations. Note that what we call 'structure' here is often called 'rule' in formal theories of grammar. Examples of *linguistic structures*, include:

- (1) a.  $S' \rightarrow (C) S$   
b. "A subordinate clause is composed of an optional complementizer and a clause".
- (2) a.  $NegP \rightarrow Adv Neg'$   
 $Neg' \rightarrow Neg VP$   
b. "Negative structures are phrases with a negative adverbial as specifier, a negative particle as head and a verb phrase as complement".

In contrast, the term 'rule' is reserved in this chapter for the sense denoted under (b) above, namely as *pedagogical rule*, to refer to the tool rather than to the object of instruction. A pedagogical rule then, is a metalinguistic description of the explicit cognitive procedure which the learner has to follow in order to correctly produce the target *structure*.<sup>2</sup> An example of a pedagogical rule is given in (3).

- (3) "To form the present conditional in French, take the root of the simple future verb form and add the endings of the imperfective indicative to it (-ais, -ais, -ait, -ions, -iez, -aient)".

Pedagogical rules are often considered as simplistic versions of what we have called linguistic structures. Unlike structures, which are part of larger models or theories that aim at descriptive, explanatory or even predictive adequacy, pedagogical rules are primarily intended as didactic tools whose primary goal is to facilitate the language learning process. As such, pedagogical rules can perhaps best be seen as a form of explicit competence while linguistic structures are a reflection of the implicit competence of a language user or language community.

### 2.3. Complexity

Turning to the notion of *complexity*, we follow Hulstijn & De Graaff (1994) and Robinson (1996) who distinguish two dimensions of complexity: "the

complexity of the structure the rule attempts to explain, and the complexity of the explanation itself” (Robinson 1996: 32). For the purposes of our research we have chosen to integrate both dimensions as shown in Figure 1.

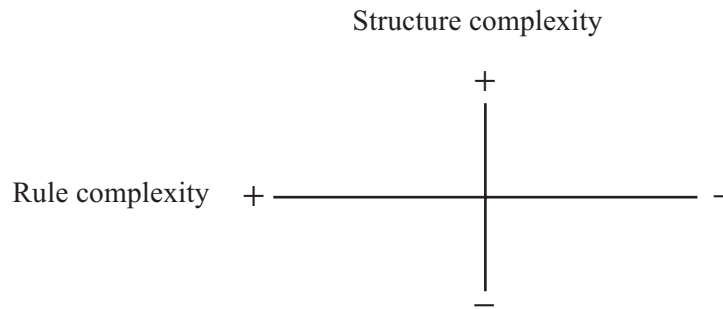


Figure 1. A model of complexity.

In our view structure complexity and rule complexity are not isomorphic: both simple and complex patterns of structural covariance can be described in both simple and complicated terms.<sup>3</sup>

### 2.3.1. Rule complexity

Py (1999) claims that one can only speak of rule complexity when out of two competing rules describing a particular grammatical phenomenon, one has a more restricted or less general quality than the other; that is, when one rule has a wider range of application in the language. It follows that complexity will be associated with the more specific rule: complex rules are rules with a relatively limited scope of application in that they apply to only a minority of items in a particular grammatical domain. In this sense, rule complexity is closely related to what Hulstijn & De Graaff (1994) call ‘rule reliability’. Py uses the conjugation of regular and irregular verbs as an example. Irregular verbs are verbs whose conjugation does not follow the general, statistically dominant pattern of morphological alternation or, in other words, verbs to which the default morphological rule does not apply. Thus, the past form of the verb ‘go’ is not ‘go+ed’ but ‘went’. In this sense, Py argues, the rule for the conjugation of the verb ‘go’ is more complex than the conjugation of the verb ‘walk’. Robinson (1996) also links the notion of complexity to the range of application of the rule but arrives at the opposite conclusion: in his opinion a simple rule is a rule which is only applicable in a limited context.

We propose to define the complexity of a pedagogical rule in terms of the *degree of elaboration* with which the rule is formulated, i.e. as the number of steps the learner has to follow to arrive at the production of the intended linguistic structure, and the number of options and alternatives available at each step. Using these parameters we can both determine the complexity of a pedagogical rule and measure the difference between the complexity levels of two competing rules. To illustrate this, consider the formation of the French present conditional. This structure can be described either by a simple rule consisting of two steps as in (4), or by a more elaborate, complex rule consisting of four steps as in (5).

(4) simple rule:

1. take the stem of the *future simple* form of the verb;
2. add the endings of the *imparfait* in the corresponding person and number (*-ais, -ais, -ait, -ions, -iez, -aient*).

(5) complex rule:

1. determine the verb class to which the verb belongs;
2. if the verb belongs to the *-er/-ir* class, then select the infinitive *être* and the endings of the *imparfait* (*-ais, -ais, -ait, -ions, -iez, -aient*);
3. if the verb belongs to the *-re* class: select the infinitive + the endings of the *imparfait*;
4. if the verb belongs to neither of the categories above, select the stem of the *futur simple* and add the endings of the *imparfait*.

This approach allows for competing rules to be hierarchically ordered in an objective manner in terms of complexity.

### 2.3.2. Structure complexity

In SLA research the expression ‘complexity of linguistic structures’ has been interpreted in at least two different ways. First, in a psycholinguistic sense, as ‘hard or difficult to process, acquire, or verbalize’ and, secondly, in a linguistic or structural sense, as ‘consisting of different components’, ‘involving derivations from a base structure’, ‘being marked or peripheral rather than default or core’ (cf. Dietz 2002). According to DeKeyser (1998) and Doughty & Williams (1998), linguistic structures in SLA research are generally classified by their linguistic rather than their psycholinguistic complexity. A further distinction is sometimes made between the *formal*

and *functional complexity* of a linguistic structure, which together determine its *overall linguistic complexity* (Ellis 1997; DeKeyser 1998; Doughty & Williams 1998b). Functional complexity refers to the degree of multiplicity (or transparency) of the mapping between the form and function of a structure. Univocal structures with a clear isomorphy between form and function (1 form – 1 function) are said to be simpler than plurivocal structures where there is no such isomorphy (1 form – n functions; n forms – 1 function). The concept of formal complexity is exemplified in Doughty & Williams (1998b) by the degree of elaboration of a morphological paradigm (e.g. the elaborate Slavic vs. the simple English present tense verb paradigm) and the dependency distance between grammatical markers (e.g. simple local inflectional morphology like tense or number marking vs. complex global inflectional morphology like subject-verb marking). According to Doughty and Williams (1998b), structures that are formally and/or functionally complex are acquired late in SLA, implying a co-occurrence of linguistic and psycholinguistic complexity.

In spite of the clear-cut theoretical demarcations between various types of structure complexity above, their operationalization in SLA research remains problematic. Different studies use different criteria to distinguish between simple and complex structures. For instance, Krashen (1982) considers the 3rd person simple present ‘-s’ marker in English as a formally simple structure because of its paradigmatic uniqueness while Ellis (1990) classifies it as formally complex because of the distance between the verb stem and the noun phrase with which it agrees. Both authors agree, however, that ‘-s’ is a functionally simple structure. In contrast, DeKeyser (1998) considers ‘-s’ to be functionally complex because of its highly syncretic nature, expressing several abstract grammatical functions simultaneously (present time, 3rd person, singular number). De Graaff (1997) operationalizes structure complexity as the total number of formal and functional grammatical criteria or features which determine the specific form and function of a given structure and which are essential for its effective noticing and processing. Yet another approach is exemplified by Robinson’s (1996) study, where expert SLA teachers were asked to identify from a list of grammatical structures the ones they thought to be more difficult for their students.

Pending a generally accepted metric for distinguishing between simple and complex linguistic structures we have, for the purposes of our study, resorted to Givon’s (1991, 1995) model of *functional markedness*. According to Givon, the functional markedness of a linguistic structure is determined by three factors: its structural complexity, its frequency and distribution, and its psycho-cognitive complexity.

- 1) structural complexity: marked structures are structurally more complex than their corresponding unmarked or less-marked structures. Structural complexity is determined by the amount of ‘linguistic substance’ of a structure (e.g. number of syllables, morphs/morphemes, phrases, clauses) or by the number of transformations from an underlying base form.
- 2) frequency and distribution: marked structures occur less frequently (esp. in general language use) and have a narrower distribution (as determined by semantic, pragmatic or grammatical constraints) than their unmarked or less-marked counterparts.
- 3) psycho-cognitive complexity: marked structures are mentally more taxing than unmarked or less-marked structures in that their production and comprehension require more and/or higher-order mental resources (e.g. attention, processing time, inductive abilities). Marked structures also develop later in language acquisition.

Our choice of Givon’s model of functional markedness was mainly motivated by practical concerns. Alternative, and perhaps theoretically more sophisticated models of markedness exist (e.g. see White 1989) but Givon’s model served our purposes best as it provides a broad descriptive framework for the selection of target structures for our experimental study of the impact of structure complexity on the effectiveness of explicit instruction with Dutch-speaking learners of French L2.

The next section describes the specific research questions and the methodology used in this study.

### **3. Study**

The general aim of our research is to investigate the role of explicit instruction in promoting L2 grammar development in an instructed L2 learning context. Specifically, we examine possible differential roles for explicit instruction depending on the complexity of the linguistic structure and of the pedagogical rules, as well as the interaction between the two. The study reported in this chapter sought to isolate the effects of the first component, the complexity of the linguistic structure, on the effectiveness of explicit instruction in promoting French-Foreign Language learners’ mastery of two French grammatical structures with different degrees of complexity. Both structures were explicitly taught to French-FL pupils receiving Dutch-medium secondary education in Flanders (Belgium).

### 3.1. Target structures

Using Givon's (1991, 1995) model of functional markedness as a heuristic, the following two grammatical structures were selected as the target of the instructional treatment: the French passive voice as the more complex/more marked structure, and French sentence negation with *ne pas*, *ne rien*, *ne personne*, *ne jamais*, *ne nulle part* as the less complex/less marked structure. These two structures differ in terms of the three criteria put forward by Givon as follows:

- 1) Structural complexity: The French passive clause is the more marked structure of the two because it involves more transformations of its underlying base form (the active clause), including:
  - modification of the verb phrase, including the addition of an auxiliary element (*être*).
  - movement of the grammatical Object to the Subject position.
  - in some cases, 'recycling' of the Subject of the active clause as the Agent of the passive clause, which involves further movement and the addition of the preposition *par*.

(6) *Le chien mord le garçon Le garçon est mordu par le chien.*

In comparison, sentence negation in French is structurally relatively simple. It merely involves the insertion of two particles in more or less fixed slots in the affirmative base structure. There is no syntactic movement or morphological modification of the verb phrase.

(7) *Il joue au football Il ne joue pas au football.*

- 2) Frequency and distribution: corpus studies suggest that passives occur far less frequently than negatives, especially in more informal spoken French, the staple input of the learners in our study (e.g. Gaatone 1998). In addition, use of the passive is subject to many more semantic and syntactic constraints than is the use of the negative. All affirmative clauses can be made negative, but not all active clauses have a passive counterpart (e.g. intransitive clauses: *Mon ami est parti en vacances hier soir*).
- 3) Psycho-cognitive complexity: We consider passives to be cognitively or psycholinguistically more complex on the basis of findings of research

on French L1 acquisition. This research has shown among other things that negatives occur from 20 months onwards whereas passives occur much later, from 48 months onwards (cf. Sabeau-Jouannet 1975). The comprehension of passives also causes problems for a much longer time than does the comprehension of negatives (e.g. Beaudichon & Lemaire 1972; Montangero 1971; Sinclair & Ferreiro 1970). Such findings suggest that the processing and acquisition of French passive clauses require more or higher-level mental resources than the processing and acquisition of French negative clauses.

### 3.2. Subjects

Participants in the study were sixty-nine Dutch-speaking secondary school pupils enrolled in three intact Grade 9 classes (age 14+) from three different secondary schools in Flanders. At the time of the study, these pupils had had French-foreign language instruction for four years (two at primary school, two at secondary school) for on average three lessons per week. According to their French teachers, the pupils had not yet been explicitly taught either target structure at the time of the study though they had obviously encountered instances of passives and especially of negatives in other French language classroom activities.

### 3.3. Design

This study is a quasi-experimental, pretest/posttest, control group design. It is quasi-experimental because the subjects were not randomly assigned to groups but rather belonged to whole classes. The three classes were thus assigned to one of two experimental groups and one control group: the first experimental group received instruction targeting the marked, complex structure (the passive) whereas the second experimental group was taught the simpler, less-marked structure (negation). The third class was assigned as a control group. The control group did not receive any special instruction but followed the regular French-Foreign Language lessons (where no instruction on the passive or negation was provided during the time of the experiment).

The design further consisted of four phases: a pretest, an instructional treatment phase, an immediate posttest and a delayed posttest (8 weeks after the first posttest).

### 3.4. Materials

#### 3.4.1. Instructional treatment

For four consecutive weeks both experimental groups received the same type of instructional treatment, which consisted of four twenty-minute grammar lessons for each target structure. These lessons followed the same scenario, consisting of the following five steps:

- 1) presentation and explanation of a metalinguistic pedagogical rule.
- 2) reading of a text containing exemplars of the target form.
- 3) identification of instances of the target form in the text.
- 4) description of these examples by referring to the pedagogical rule.
- 5) controlled practice and exercises (e.g. sentence transformation, answering semi-open questions).

The instruction in this study constituted a deliberate focus on one specific feature of the target language itself (i.e., structure) rather than an exposure to the linguistic target feature through the target language (i.e., comprehensible input). The presentation varied between inductive and deductive strategies, the practice between production and comprehension/recognition and between discrete sentences and longer segments of discourse, while the entire instructional schema attempted to engage the subjects' cognitive skills in understanding and applying the rules of the formation and use of the target structure in a systematic manner.

The instructional treatment was given in French. Separate instructional packets for both experimental groups were designed and balanced in regard to: (a) vocabulary, (b) total amount of instruction time versus practice time, and (c) total number of practice activities. None of the groups were assigned homework on the target structures during the period of the experiment, and no review of the target structures was provided during the two posttests.

#### 3.4.2. Tests

The effect of the instructional treatment was assessed by comparing learners' pretest-posttest performance on three tasks: a grammaticality judgement task, a controlled written production task and a semi-guided oral production task. Cronbach alpha values varied from .83 to .92, indicating that the tests and measures used in this study can be considered as reliable.



In the grammaticality judgement task the learners had to evaluate on a 3-item scale (correct – no opinion – incorrect) a set of sixty sentences containing grammatical and ungrammatical instances of the target structure and rule.

(8a) *Les résultats sont donnés par le professeur.* ('The results are given by the professor')

(8b) *Je pars ne pas demain.* ('I leave not tomorrow')

The controlled written production task consisted of eighteen semi-open questions asking the learners to turn active and affirmative sentences into respectively passive and negative sentences (cf. examples 9a/b):

(9) *Turn the following sentence to passive/negative:*

(a) *Les gardes du parc protègent les éléphants.* > *Les éléphants sont protégés par les gardes du parc.* ('The zoo-keepers protect the elephants' > 'The elephants are protected by the zoo-keepers.')

(b) *J'ai vu quelqu'un dans la maison.* > *Je n'ai vu personne dans la maison.* ('I have seen someone in the house.' > 'I haven't seen anyone in the house.')

Learners' oral production was assessed during an oral interview with open questions designed to elicit unplanned production of the target structures. In the oral interview, learners were asked open questions about pictures and objects, which they had to answer instantly (so as to leave no time for planning; cf. below).

(10a) *Pourquoi ce monsieur est-il triste* → *Parce qu'il n'a pas d'argent*  
(Why is that man sad? → Because he doesn't have any money)

(10b) *Qu'est-ce qui se passe avec les personnages en bleu?* → *La fille est embrassée par un garçon*  
(What is happening to the characters in blue? → The girl is being kissed by a boy.)

The three tasks in this study are assumed to invoke different types of language knowledge and processing mechanisms. Previous studies on the effects of instruction (e.g. Fotos & Ellis 1991; Fotos 1993, 1994; Van Patten & Cadierno 1993) typically measured the change induced by instruction on discrete item tasks only (e.g. fill in the gap exercises, sentence transformation

tasks). Such tasks involve primarily explicit, memory-based performance. These studies have shown that language instruction can help controlled production of the target structures. But there is no clear evidence yet that explicit instruction also results in either short or long term communicative use of the target structure (see reviews in Norris & Ortega 2000; Ellis 2002). Inspired by information processing theories (e.g. Anderson 1993), 'effects of instruction' studies increasingly opt for the inclusion of multiple tasks (e.g. White 1998) and investigate the effects of planning (e.g. Skehan & Foster-Cohen 1996; Skehan & Foster 1999) to assess the various types of knowledge used by language learners. The assumption made in this study, then, is that the three tasks used differ in their degree of planning and therefore in the activation of explicit vs. implicit knowledge. The grammaticality judgement task and the controlled written production task provide the learners with planning time and therefore draw primarily on explicit knowledge. The oral open interview task was included on the recommendation of Ellis (2001, 2002) and others to determine whether the learners in the experimental groups were able to apply the rules they were taught in their unplanned language use. This task was so construed that it offered little or no planning opportunity, thus limiting the activation of explicit knowledge and forcing the learners to rely more on the activation of implicit knowledge.

### 3.5. Procedures

#### 3.5.1. *Treatment and data collection*

For the sake of the ecological validity of the study and its results, and following previous research (e.g. VanPatten & Cadierno 1993), all instruction and assessment took place in the pupils' regular classrooms during their normal class times. In order to ensure the standardization of the experimental conditions, one researcher administered the instructional treatment to the two experimental groups and also administered all the tests.

#### 3.5.2. *Scoring procedures*

The grammaticality judgement task in each of the three tests was worth 60 points. One point was assigned for each correct response; incorrect responses and 'no opinion' responses received a score of zero.

The controlled written production task was worth a total of 18 points. Again, one point was given for each correct target structure provided. Incorrect structures, or failure to produce the target structure received a zero.

The learners' utterances produced during the oral interview were all transcribed in CHAT format (MacWhinney 1995) and analysed for two aspects: the number of target forms produced and the accuracy with which these were used. We wanted to know whether the instruction would lead learners to use the target structure more often, and if so, whether they would use these forms more accurately. Accuracy was measured by a derivative of the Target-Like Usage (TLU) index (cf. Pica 1984).<sup>4</sup>

### *3.5.3. Statistical procedures*

Given the relatively small sample size and the danger of the data not being normally distributed, non-parametric statistics (Kruskal-Wallis Anova for three groups, Mann Whitney U for two independent samples and Wilcoxon for two related samples) were used to determine whether the instructional treatment had an effect on the test performance of the experimental groups and the control group and if so, whether this effect was moderated by the complexity/markedness of the target structure. The following levels of statistical significance are used: \* =  $p < .05$ ; \*\* =  $p < .01$ ; \*\*\* =  $p < .001$ .

Statistical analysis of the pretest scores revealed significant differences between the groups on several of the tasks and measures used (see 'Results' section below). Therefore it was decided to perform the analysis of pretest-posttest differences on 'gain scores'; that is, we compared the progress or regression of each group between the different test moments and determined whether the difference between the gains of each group was significant or not.<sup>5</sup>

## **4. Results**

As a preliminary research question we will first examine whether the instructional treatment had an effect on the pupils' mastery of the target structures (section 4.1). Then we will examine whether this effect is moderated by the complexity/markedness of the target structure (section 4.2).

### **4.1. Does explicit instruction have an effect?**

In order to answer the first research question, the scores of the two experimental instruction groups were combined (+Instruction Group) and compared to the scores of the control group (-Instruction Group). The charts in

figures 2-5 show the evolution of the test scores over time for each of the three tasks. They reveal a significant difference on the pretests between the +Instruction group and the -Instruction group on the grammaticality judgement task. Tables 1–4 show the results of the corresponding gain score analysis. (A positive difference in the third column of these tables indicates that the gain was greater for the +Instruction group than for the -Instruction group, a negative difference indicates that the gain was greater for the Control group than for the +Instruction group).

#### 4.1.1. Grammaticality judgement task

Figure 2 and table 1 show that the Instruction group obtained significantly higher scores on the grammaticality judgement task than the control group on all three tests. The scores of both groups increase immediately after instruction (by 6.1% for the +Instruction group and by 3.4% for the -Instruction group) but this initial gain does not appear to be durable as the scores of both groups decrease again in the eight week period between the first and the second posttest. Table 1 further shows that the gains of the +Instruction group are consistently higher than those of the control group. However, the differences between the gain scores of the two groups are not significant. This means that the treatment (explicit instruction) had no significant effect on subjects' performance on the grammaticality judgement task.

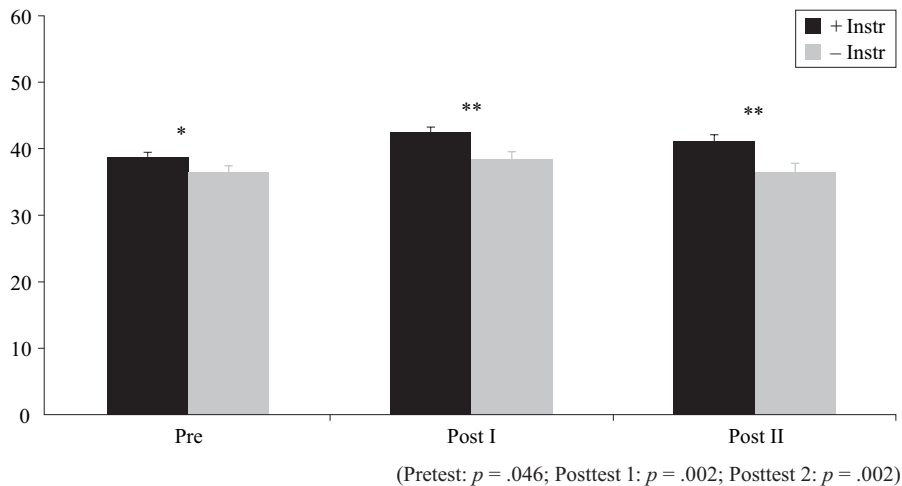


Figure 2. Comparison between the scores of the +Instruction Group and -Instruction Group: grammaticality judgement task

Table 1. Gain scores of the +Instruction Group and –Instruction Group: grammaticality judgement task.

	+ Instruction		– Instruction		Difference		
	score	%	score	%	score	%	<i>p</i>
Pre-Post1	3.64	6.1	2.05	3.4	1.59	2.7	.154
Pre-Post2	2.36	3.9	0.08	0.1	2.28	3.8	.072
Post1-Post2	-1.29	-2.1	-1.98	-3.3	0.69	1.1	.444

4.1.2. Controlled written production task

Figure 3 and table 2 show that the +Instruction group again significantly outperforms the –Instruction group on the written production tasks on all three tests. The scores of both groups also progress from the pretest to the first posttest. In the case of the +Instruction group the initial gain appears to be durable as its posttest 2 scores are even higher so that its final net gain amounts to nearly 20 percentage points when compared to the pretest score. In contrast, the scores of the control group on this tasks decrease by nearly 2% in the eight week period between the first and second posttest, resulting in a final gain of 6.3%. Comparison of the gain scores of the two groups revealed significant differences (see table 3), indicating an effect of the instructional treatment on subjects’ performance on the written production task.

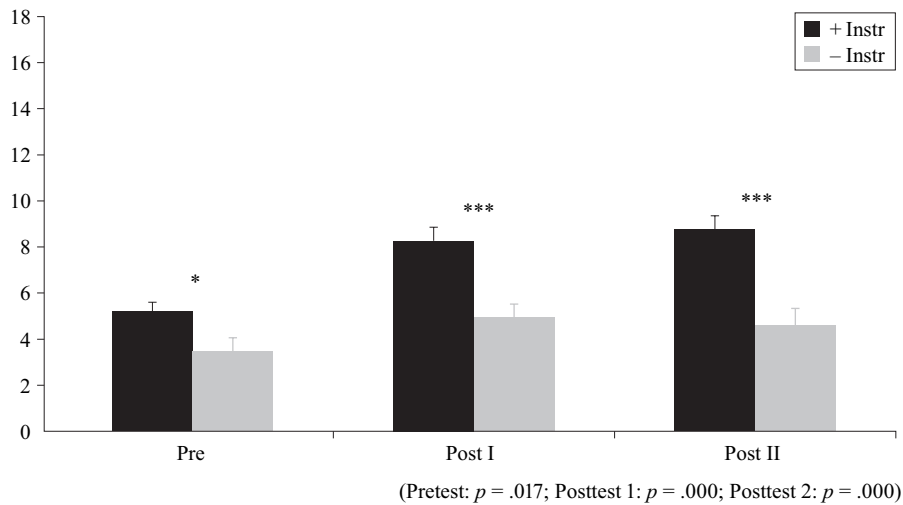


Figure 3. Comparison between the scores of the +Instruction Group and –Instruction Group: controlled written production task.

Table 2. Gain scores of the +Instruction Group and –Instruction Group: controlled written production task

	+ Instruction		– Instruction		Difference		<i>p</i>
	score	%	score	%	score	%	
Pre-Post1	3.08	17.1	1.48	8.2	1.61	8.9*	.012
Pre-Post2	3.56	19.8	1.13	6.3	2.44	13.5**	.001
Post1-Post2	0.48	2.7	–0.35	–1.9	0.83	4.6	.285

#### 4.1.3. Unplanned oral production task

As for the oral production task, figures 4 and 5 show that both groups advance after treatment both for the general productivity measure and the accuracy measure, but the instructed group clearly more so than the control group. Note that the control group outperformed the experimental group on both measures on the pretest, but not significantly so. When we look at the posttest scores, we see that the experimental group outperforms the control group, again on both measures of oral production. However, the differences between the two groups are significant only for the accuracy measure (figure 5), not for the general productivity measure (figure 4).

Turning to the results of the gain score analysis, we see in table 3 that the +Instruction group produced around 30 tokens more of the target structure after instruction than before instruction. The control group produced only about five tokens more. The differences between the gains of the two groups are significant. Table 4 demonstrates that during the oral interviews the instructed group became 21% more accurate in their use of the target structures after treatment. The accuracy of the control group increased as well but less strongly so (5% on posttest 1, 11% on posttest 2). The differences between the gains of the two groups are again significant.

Table 3. Gain scores of the +Instruction Group and –Instruction Group: general oral productivity (*absolute gain in number of tokens produced*)

	+ Instruction	– Instruction	Difference	<i>p</i>
	score	score	score	
Pre-Post1	27.96	5.18	22.78***	.000
Pre-Post2	34.84	12.50	22.34***	.001
Post1-Post2	6.88	7.33	–0.45	.860

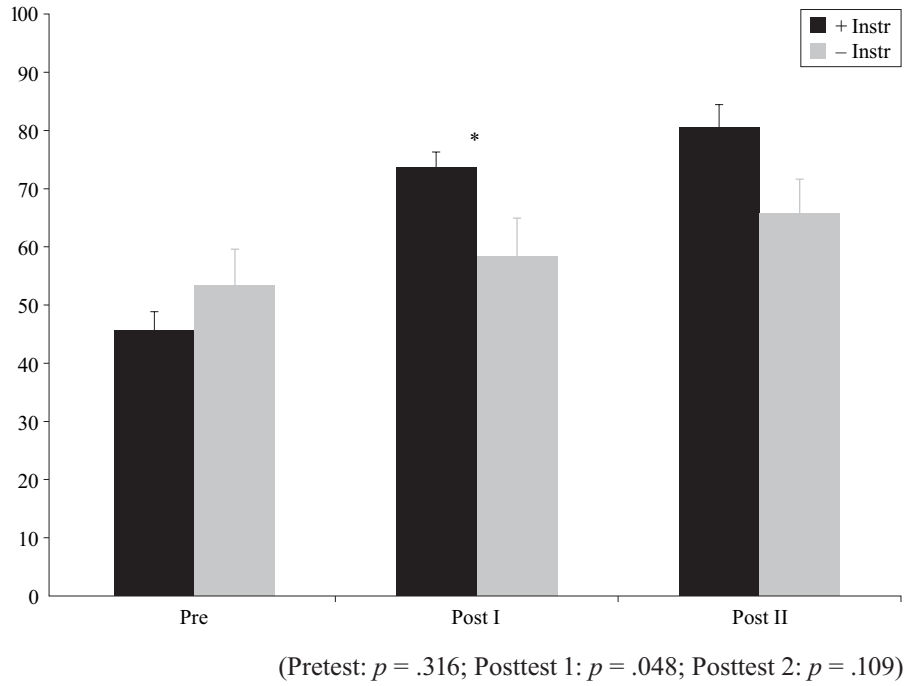


Figure 4. Comparison between the scores of the +Instruction Group and -Instruction Group: general oral productivity

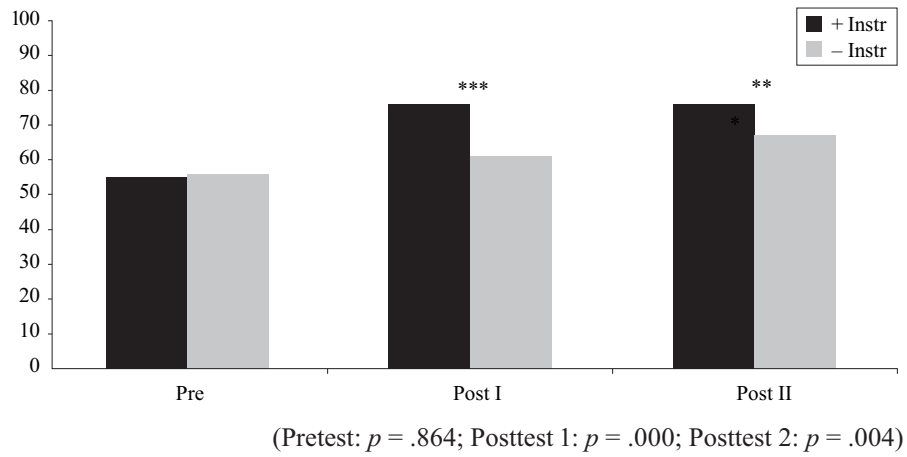


Figure 5. Comparison between the scores of the +Instruction Group and -Instruction Group: accuracy of oral production

Table 4. Gain scores of the +Instruction Group and -Instruction Group: accuracy of oral production (*as percentage*)

	+ Instruction %	- Instruction %	Difference %	<i>p</i>
Pre-Post1	21.0	5.0	16.0***	.008
Pre-Post2	21.0	11.0	10.0*	.121
Post1-Post2	0.0	6.0	-6.0	.228

#### 4.1.4. Summary

Since the experimental group advanced significantly more than the control group between pretest and posttest scores on the written and oral production tasks – though not on the grammaticality judgement task – we can conclude that the type of explicit instruction employed in our study had a beneficial effect on the subjects' mastery of the grammatical target structures, or at least on their productive mastery. This finding is in line with previous research reviewed in Norris and Ortega (2000) and Ellis (2001) which found that “FFI [form-focused instruction], *especially of the more explicit kind*, is effective in promoting language learning” (Ellis 2001: 12; italics added). What is interesting about our results is that the instructional effect is also observed, and even most strongly so for the task that draws most on implicit knowledge, i.e. the unplanned oral production task.

#### 4.2. Impact of structure complexity on the effectiveness of instruction

Having established that the instruction indeed has an effect on learners' mastery of the target structure, we turn to the second research question and examine whether this effect is modified by the markedness/complexity of the target structure: “Does the instruction have a greater impact on the learning of the complex structure than on the learning of the simpler structure (or vice versa)?”

To investigate this question, we compared the scores of the group which received instruction on the complex structure (the passive) with the scores of the group which was taught the simpler structure (negation). In what follows, the former group will be referred to as the “Complex Group”, the latter as the “Simple Group”. The charts in figures 6–9 again illustrate the



development of the scores on the three tests for each of the three tasks. Tables 5–8 show the results of the corresponding gain score analyses. (A positive difference in the third column of these tables indicates that the gain was greater for the Complex group than for the Simple group, a negative difference indicates that the gain was greater for the Simple group than for the Complex group).

#### 4.2.1. Grammaticality judgement task

Figure 6 shows that the Simple group consistently and significantly outperforms the Complex group on this task. This result simply means that these learners are better at making grammaticality judgements about negative structures than about the passive structures, before as well as after the instructional treatment. This result does not necessarily mean, however, that the instruction had a greater impact on learners' ability to judge simple structures than on their ability to judge complex structures. In order to investigate that possibility, we have to look at table 5 and compare the differences in the gains made by the two groups after instruction.

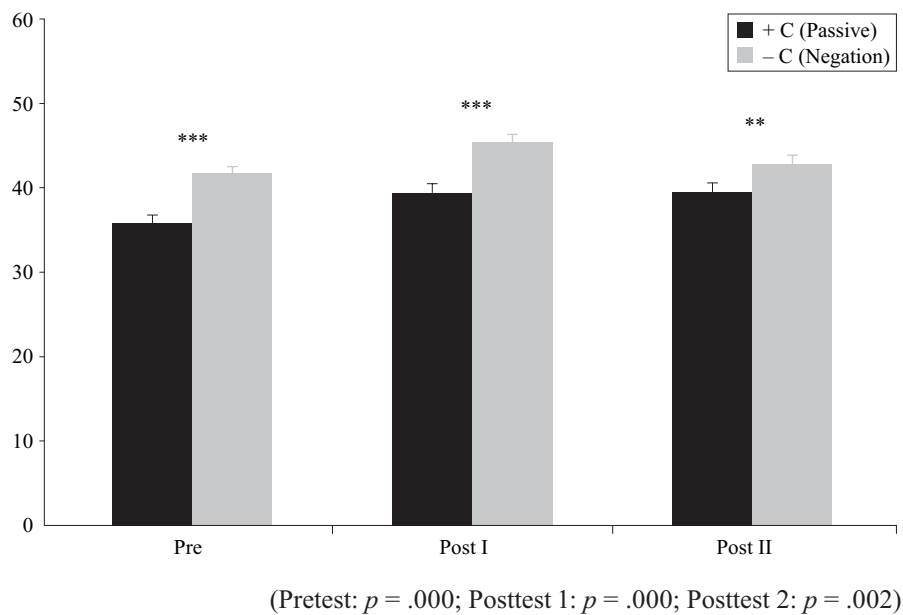


Figure 6. Comparison between the scores of the Simple Group (-C) and the Complex Group (+C): grammaticality judgement task.

Table 5 shows that the Complex group gained 5.9% on posttest 1, which slightly increased to 6% on posttest 2, while the Simple group gained 6.2% and 1.9% respectively. The differences between these gains are not significant. When we compare the gains between the two posttests we see that the Simple group loses much of the gain made immediately after treatment (−4.3%), whereas the initial gain of the Complex group is sustained (+0.1%). The difference between these gain scores is significant. Thus, although both groups make similar gains immediately after instruction, the gain of the Simple group is not durable. This result could indicate that the complexity of the target structure does not necessarily influence the effect of explicit instruction on learners' ability to make grammatical judgements, but rather, that whatever gains are made in this regard from instruction, they tend to diminish faster in the case of simpler structures.

Table 5. Gain scores of the Simple Group and the Complex Group: grammaticality judgement task

	Complex		Simple		Difference		<i>p</i>
	score	%	score	%	score	%	
Pre-Post1	3.55	5.9	3.73	6.2	−0.18	−0.3	.758
Pre-Post2	3.59	6.0	1.12	1.9	2.47	4.1	.082
Post1-Post2	0.04	0.1	−2.61	−4.3	2.65	4.4*	.013

#### 4.2.2. Written production task

The picture that emerges from figure 7 and table 6 for the written production task is similar to that of the grammaticality judgement task described above. Figure 7 indicates that the performance of the Simple group on this task is significantly better than the performance of the Complex group for all three tests. (As such this is not a very interesting result though it provides additional confirmation of our assumption that the French passive is indeed more marked/complex than French negation). Both groups also perform substantially better on the written production task after explicit instruction than before. The greatest gain in scores is made immediately after the treatment (posttest 1). Table 6 shows that the Complex group gains most (with a 23.6% gain on posttest 2 as opposed to a 16.1% gain for the Simple group) but the differences in gain between the two groups are not significant. Again we must conclude that the instruction had a similar effect on learners'

controlled written production of both target structures, regardless of their complexity.

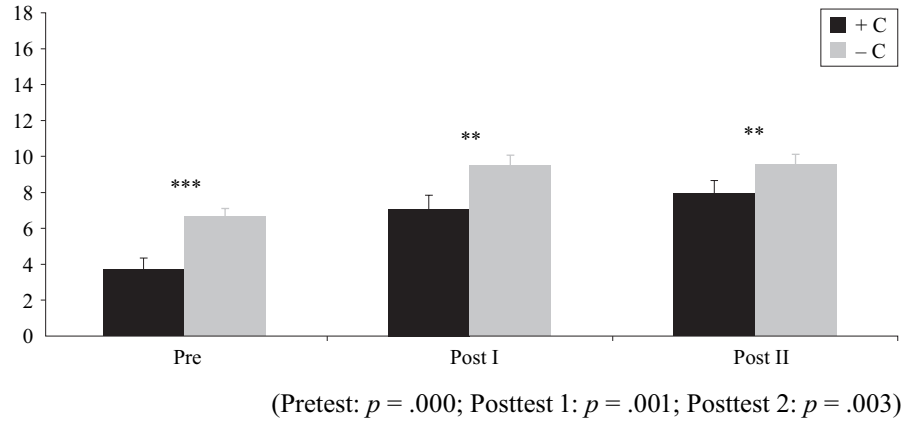


Figure 7. Comparison between the scores of the Simple Group (-C) and the Complex Group (+C): controlled written production task.

Table 6. Gain scores of the Simple Group and the Complex Group: controlled written production task

	Complex		Simple		Difference		
	score	%	score	%	score	%	<i>p</i>
Pre-Post1	3.37	18.7	2.80	15.6	0.57	3.1	.364
Pre-Post2	4.24	23.6	2.88	16.1	1.36	7.5	.134
Post1-Post2	0.87	4.9	0.08	0.5	0.79	4.4	.113

#### 4.2.3. Oral production

Next we consider the learners' unplanned oral production skills, starting with the frequency with which the two instructed groups use their respective target structure before and after instruction.

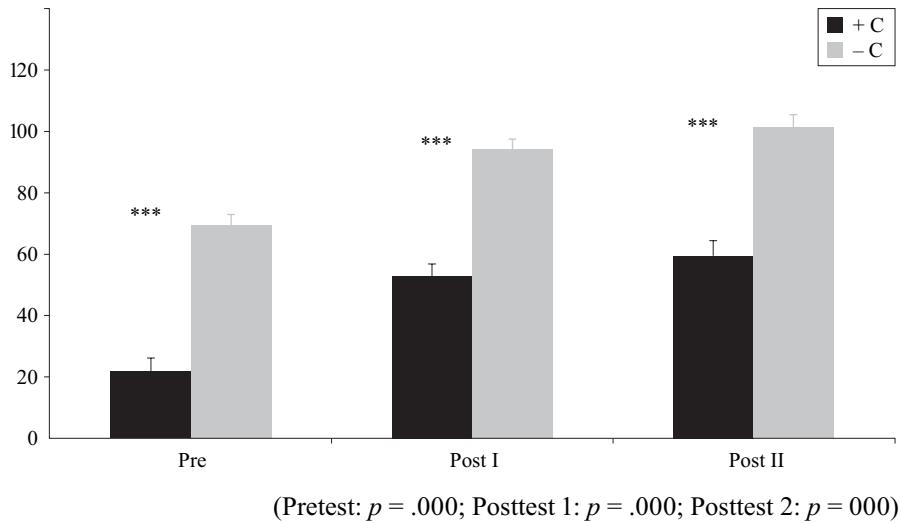


Figure 8. Comparison between the scores of the Simple Group (-C) and the Complex Group (+C): general oral productivity

Table 7. Gain scores of the Simple Group and the Complex Group: general oral productivity (absolute gain in number of tokens produced)

	Complex %	Simple %	Difference %	$p$
Pre-Post1	31.12	24.80	6.32	.334
Pre-Post2	37.84	31.84	6.00	.451
Post1-Post2	6.72	7.04	-0.32	.675

Figure 8 first shows that the Simple group produces overall many, and significantly more negatives than the Complex group produces passives, before as well as after instruction. Probably more interesting is the fact that both groups produce significantly more instances of their respective target structures after the instruction than before, particularly *immediately* after the instruction (i.e. on posttest 1). Table 7 further shows that the Complex group gains most (some 38 tokens more on posttest-2 than on the pretest, as opposed to 32 tokens more for the Simple group). However, the difference in progress is once more not significant, suggesting that the effect of explicit instruction on this measure of unplanned production is not modified by the complexity of the target structure.

Finally, we turn to the accuracy with which the instructed subjects use their respective target structure in their unplanned speech.

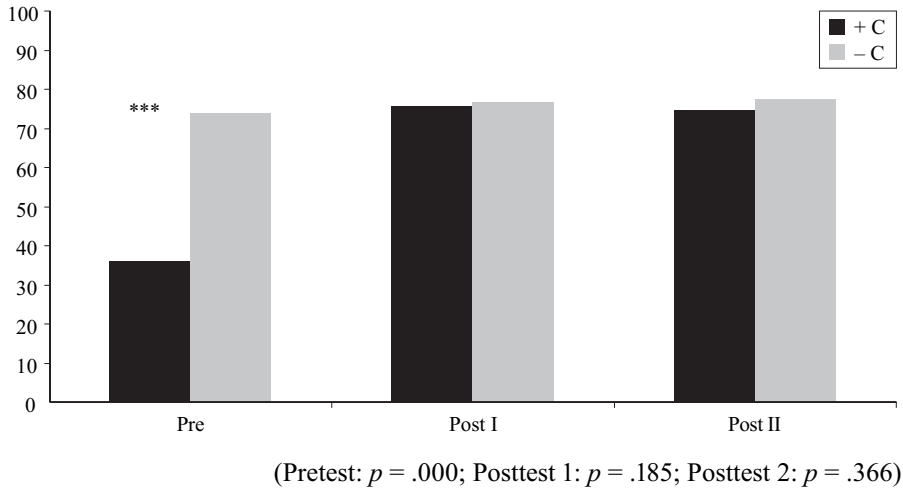


Figure 9. Comparison between the scores of the Simple Group (-C) and the Complex Group (+C): accuracy of oral production.

Table 8. Gain scores of the Simple Group and the Complex Group: accuracy of oral production (as percentage)

	Complex %	Simple %	Difference %	<i>p</i>
Pre-Post1	39.7	2.7	37.0***	.000
Pre-Post2	38.6	3.5	35.1***	.000
Post1-Post2	-1.1	0.8	-1.9	.946

Figure 9 shows that with a TLU score well below 40%, the Complex group made many errors in their production of the passive during the pretest – and significantly more so than the Simple group – but they also greatly improved after the instructional treatment, reaching TLU accuracy scores similar to that of the Simple group on the posttests. No such strong instructional effect is observed for the Simple group, whose accuracy scores increase only minimally. Table 8 shows that the differences between the gain scores of the two groups are significant. This could indicate that the complexity of the target structure has an effect on the efficacy of the instruction. Never-

theless, we hesitate to conclude from this result that explicit instruction has a greater impact on the accuracy of the complex structure than on the accuracy of the simple structure, for two related reasons. First, the scores of the Complex and Simple groups on the accuracy measure already differed significantly at the time of the pretest (Complex group=36%, Simple group=74%). While this is true for several of the other pretest scores discussed so far, the pretest difference is never as great as on this measure (38 percentage point difference). Second, with a TLU pretest score of 74%, the Simple group already had a high, close to native speaker accuracy level (conventionally set between 90–100%) before the instruction, leaving this group little room for improvement. In order to resolve the first problem, we chose to compare gain scores rather than absolute scores. However, the use of gain scores is not unproblematic: a gain score analysis assumes that skill development proceeds in a linear manner whereas, in fact, several aspects of language proficiency probably develop *asymptotically*, with development gradually trailing off as it approaches a natural plateau or ceiling level. This is a general problem not only with gain scores but also with many index scores, such as the TLU, when used as measures of skill or knowledge development. Their use is based on the assumption that an increase from 20 to 30% amounts to the same as an increase from 90 and 100%. This assumption may not hold for certain features of language.

#### 4.2.4. Summary

Examination of the gain scores of the Complex and the Simple group after the instruction reveals that the greatest gains from the pretest to the post-tests are nearly always made by the group which was taught the complex structure (see the positive values in the third column in tables 5–8). These differences in gains, while suggestive of an effect of the complexity/ markedness of the target structure on the efficacy of the instruction, are rarely statistically significant. We must therefore conclude that the target structure had a nugatory influence on the effectiveness of explicit instruction in promoting subjects' ability to make grammaticality judgements about their respective target structures, on their ability to use their target structure in a controlled written production task, and on the frequency with which they produce the target structure in a semi-guided oral interview. Statistically significant differences between the two groups were observed in their progress in oral accuracy but caution is warranted in attributing these differences to the difference in markedness/complexity between the two target

structures, given the magnitude of the pretest difference between the two groups as well as the limitations of the accuracy measure used in this study.

## **5. Conclusion**

The general question addressed in the study reported in this chapter is: “Does the nature of a linguistic structure determine the extent to which it can be learned through explicit grammar instruction?” Or, in even more general terms, “Are some linguistic structures more teachable than others?” Claims to this effect have frequently been made in the SLA literature but insufficiently demonstrated. In this study we hypothesized that explicit instruction would affect French-FL learners’ mastery of two grammatical target structures, French passive and negation, resulting in increased performance on a variety of language tasks (involving both receptive and productive skills and both explicit and implicit language knowledge). We further hypothesized that the different degrees of complexity of the target structures, here defined in terms of their functional markedness (Givon 1995), would differentially affect the effectiveness of the instructional treatment, resulting in greater gains in the case of one of the two structures.

The results of the analyses lead us towards two tentative conclusions. First, they suggest a beneficial effect of explicit instruction on learners’ mastery of two different grammatical structures, at least when we consider their productive language use. The fact that this beneficial effect is even more strongly observed in the subjects’ unplanned speech, as elicited in a semi-guided oral interview with little or no time for planning, could be taken to suggest that explicit instruction not only promotes explicit grammatical knowledge, as shown by previous studies, but also implicit knowledge. The exact nature of the underlying mechanisms responsible for this process are still unclear, and cannot be deduced from the results presented here. One possibility is that the explicit instruction affected the salience of the target structures by directing the learners’ attention toward them, thus leading to increased awareness and noticing (cf. Schmidt 1995; Ellis & Laporte 1996).

Secondly, the scores from the group with explicit instruction on the passive (the marked/complex structure) and the group with explicit instruction on negation (the less marked/simpler structure) did not reflect any clear effect of the target structure on the efficacy of instruction. This would suggest that structural complexity is less significant than hypothesised for the

effects of explicit instruction. However, on a more positive note, learners who were taught the more marked structure systematically gained more from this than did learners who were taught the less marked structure. Again, the differences are often not significant but suggestive nevertheless of a relationship between the complexity of the target structure and the impact of explicit instruction on the L2 learning process.

Clearly, the results of this study must be interpreted with caution because of the limitations of the design and of the analytic procedures used, and also because of conceptual limitations. First, there is the problem that the experimental and control groups showed different levels of mastery of the target structures before the instructional treatment. Initial, pre-treatment differences are a risk inherent to the non-random sampling procedure which is a characteristic of the quasi-experimental design. There are several ways of dealing with such pretest differences. One method, the use of gain scores, has specific limitations, as illustrated by the present study. Therefore, future analyses may consider alternative methods, including other statistics (e.g. covariance analysis, with the pretest scores as a covariate to filter out the pretest differences between the groups).<sup>6</sup>

In a follow-up study we will also consider other, more sensitive measures and indexes of unplanned oral production than the productivity and accuracy measures used in the present study, as well as other tasks for tapping implicit knowledge to complement the semi-guided interview (e.g. narration and decision making tasks). As Ellis (2001) has pointed out, time pressure does not necessarily guarantee a measure of implicit knowledge as some learners may have developed automatized explicit knowledge which they can apply even under time pressure. Consequently, language tasks which allow little or no planning time (like the oral interview in this study) may not necessarily provide appropriate measures of learners' implicit knowledge, but rather, of procedural knowledge. Future research should attempt to develop tasks which can distinguish between proceduralized explicit knowledge and proceduralized implicit knowledge.

There is also a need for more precise ways of defining and operationalizing the notion of 'structure complexity'. Although Givon's model of functional markedness was chosen as a descriptive framework for the selection of target structures in the present study, it suffers from some conceptual vagueness and circularity. This may have resulted in the selection of target structures which were perhaps insufficiently matched.

Finally, in our own research we will refine the present analyses by including other possible mediating factors, such as the *complexity of the*



*metalinguistic pedagogical rule* used to teach the target structure. In fact, the data presented in this chapter are part of a larger study with a more complex design: for each of the two target structures we have formulated both a *simple rule* and a more complex or *elaborate rule* on the basis of the procedure outlined in section 2. Both types of rules were used for teaching both types of target structures. The results of the analysis of the effect of rule complexity and the interaction between rule complexity and structure complexity are still pending but it may well turn out that, for maximal instructional effect, a complex target structure requires an elaborate pedagogical rule, whereas a simple structure requires a simple rule. Or vice versa. Whatever the case may be, it is clear that the effects, and the effectiveness, of explicit instruction depend on a wide range of factors, and that the exact nature and role of even one such factor, such as complexity, is by no means straightforward. There is a clear need for further experimental research and for further theoretical clarification of the factors that moderate the effectiveness of explicit instruction.

## Notes

1. The authors wish to thank the following people who contributed a great deal to the experimental study reported in this chapter: Géraldine De Visscher, for designing and administering the instructional and testing materials and for analysing the data, and Geoffrey De Visscher, for the statistical analyses.
2. Our definition of ‘rule’ corresponds to that of Gor & Chernigovskaya (this volume), who define it as follows: “The term ‘rule’ refers to the pedagogical rules, or in other words, explicit explanations provided to the learner in the classroom. It is not intended to mean symbolic rules used in linguistic descriptions or psycholinguistic rules involved in mental processing”.

Our definition thus slightly deviates from that of Doughty & Williams 1998: “In our discussion .. the terms forms and rule are both used, since learners are engaged in acquiring both [...] Put simply, rules describe the realization, distribution, and use of forms. Thus, for us, both forms and rules are subsumed by the more comprehensive term *form*. Furthermore, it is important to note that every hierarchical level of language – from phonology to morphosyntax to the lexicon to discourse and pragmatics – is composed of both forms (e.g. phonemes, morphemes, lexical items, cohesive devices, and politeness markers) and rules (e.g. devoicing, allomorphy, agreement, collocation, anaphora, and in-group vs. out-group relationships)” (p. 211–212).

3. De Graaff (1997) has pointed out that a strict separation of rule complexity and structure complexity may not always be possible in practice: rules which describe structures which are formally and/or functionally complex tend to be more elaborate than rules describing simple structures.
4. The following ratio was used in the computation of the TLU index score:
 
$$\frac{(\text{correct occurrences in appropriate contexts } \times 2) + (\text{malformed occurrences in appropriate contexts})}{(\text{appropriate contexts} + \text{inappropriate occurrences}) \times 2}$$
5. The gain from the pretest to the posttests is computed for each learner and group of learners by subtracting each learner's pretest score from his posttest scores, and by subtracting the posttest-1 scores from the posttest-2 scores. When computed in this manner, a positive gain score indicates that the posttest score was greater than the pretest score, a negative gain score indicates that the posttest score was less than the pretest score. In our case the dependent variable is performance on three language tasks so we expect that successful instructional treatment would lead to better performance. The gain score should therefore be positive. The differences between the pretest and posttest scores for each subject and group were analysed in a one-way analysis of variance using instructional treatment (experimental vs. control) as the only factor for the analyses presented in section 4.1 and type of structure (simple vs. complex) as the only factor in the analysis in section 4.2. If the treatment main effect is significant, then the change from pretest to posttest is not the same in the two groups. This analysis of difference scores is called a *gain score analysis* (see Becker 1999; Gardner 1987).
6. Still another method for solving the problem of pretest differences is examining the interaction effect in a 2 x 2 analysis of variance with treatment (treatment vs. control) as a between subjects factor and time (pretest vs. posttest) as a within subjects factor. If the interaction is significant, then the change between pretest and posttest is not the same in the two treatment conditions (see Becker 1999; Gardner 1987).

## References

- Alanen, R.  
1995 Input enhancement and rule presentation in second language acquisition. In *Attention and awareness in foreign language learning and teaching*, R. Schmidt (ed.), 259–302. Honolulu: University of Hawaii Press.
- Allen, J. F., J. Hendler and A. Tate (eds.)  
1990 *Readings in Planning*. San Mateo, CA: Morgan Kaufmann Publishers.
- Anderson, J. R.  
1993 *Rules of the Mind*. Hillsdale, NJ: Erlbaum.

- Beaudichon, J. and F. Lemaire  
1972 Acquisition de la négation considérée sous différents aspects chez l'enfant de 3 à 8 ans. *Bulletin de psychologie* 26: 375–382.
- Becker, L.  
1999 *Analysis of Pretest and Posttest Scores with Gain Scores and Repeated Measures*. <http://web.uccs.edu/lbecker/Psy590/gainscore.htm>.
- Bley-Vroman, R. W.  
1989 The logical problem of second language learning. In *Linguistic Perspectives on Second Language Acquisition*, S. Gass and J. Schachter (eds.), 41–68. Cambridge: CUP
- Carroll, S. and M. Swain  
1993 Explicit and implicit negative feedback: an empirical study of the learning of linguistic generalizations. *Studies in Second Language Acquisition* 15 (3): 357–386.
- De Graaff, R.  
1997 *Differential Effects of Explicit Instruction on Second Language Acquisition*. Den Haag : Holland International Graphics.
- DeKeyser, R.  
1998 Cognitive perspectives on learning and practicing L2 grammar. In *Focus on Form in Classroom Second Language Acquisition*, Doughty, C. and J. Williams (eds.), 41–68. Cambridge: CUP.
- Dietz, G.  
2002 On rule complexity: A structural approach. *EUROSLA Yearbook August 2002*, 2 (1): 263–286.
- Doughty, C. and J. Williams  
1998 Pedagogical choices in focus on form. In *Focus on Form in Classroom Second Language Acquisition*, Doughty, C. and J. Williams (eds.), 197–261. Cambridge: CUP.
- Doughty, C. and J. Williams (eds.)  
1998 *Focus on Form in Classroom Second Language Acquisition*. Cambridge: CUP
- Ellis, N.  
2002 Frequency effects in language processing. *Studies in Second Language Acquisition* 24: 143–188.
- Ellis, N. C. and N. Laporte  
1996 Contexts of acquisition: effects of formal instruction and naturalistic exposure on second language acquisition. In *Tutorials in bilingualism: Psycholinguistic perspectives*, A. de Groot and J. Kroll (eds.), 53–83. Hillsdale, NJ: Erlbaum.
- Ellis, R.  
1990 *Instructed Second Language Acquisition*. Oxford: Basil Blackwell.  
1994 A theory of instructed second language acquisition. In *Implicit and Explicit Learning of Language*, N. Ellis (ed.), 79–114. London: Academic Press.

- 1997 *Second Language Acquisition*. Oxford: Oxford University Press.
- 2001 Investigating form-focussed instruction. *Language Learning 51: supplement 1. Form-focussed instruction and Second Language Learning*, 1–46.
- 2002 Does Form-Focused Instruction affect the acquisition of implicit knowledge? A review of the research. *Studies in Second Language Acquisition*, 24: 223–236.
- Felix, S. and W. Weigl  
 1991 Universal Grammar in the classroom: the effects of positive and negative evidence in the classroom. *Second Language Research 7 (2)*: 162–180.
- Fotos, S.  
 1993 Consciousness-raising and noticing through focus on form: Grammar task performance versus formal instruction. *Applied Linguistics 14*: 385–407.
- 1994 Integrating grammar instruction and communicative language use through grammar consciousness-raising tasks. *TESOL Quarterly 28*: 323–351.
- Fotos, S. and R. Ellis  
 1991 Communicating about grammar instruction and communicative language use through grammar consciousness-raising tasks. *TESOL Quarterly 25*: 605–628.
- Gaetone, D.  
 1998 *Le passif en français*. Louvain-la-Neuve. Paris: Duculot.
- Gardner, R. C.  
 1987 Use of the simple change score in correlational analysis. *Educational and Psychological Measurement, 4*: 849–864.
- Givón, T.  
 1991 Markedness in grammar: Distributional, communicative and cognitive correlates of syntactic structure. *Studies in Language 15 (2)*: 335–370.
- 1995 *Functionalism and Grammar*. Amsterdam-Philadelphia: John Benjamins.
- Harley, B.  
 1994 Appealing to consciousness in the second language classroom. *AILA Review 11*: 57–68.
- Hendrix, L., Housen, A. and M. Pierrard  
 2002 Le rôle de mode d'implémentation de l'instruction grammaticale dans le processus d'acquisition de langues étrangères. *Acquisition et Interaction en Langue Etrangère 16*: 73–96.
- 2003 Impact et efficacité de différents types d'enseignement axés sur la forme pour l'acquisition d'une langue étrangère. *Marges Linguistiques 4*: 23–49.

- Hulstijn, J.  
1989 A cognitive view on interlanguage variability. In *The dynamic interlanguage: empirical studies in second language acquisition*, M. R. Eisenstein (eds.). New York: Plenum Press.  
1995 Not all grammar rules are equal: Giving grammar instruction its proper place in foreign language teaching. In *Attention and awareness in foreign language learning*, R. Schmidt (ed.), 359–386. Honolulu: University of Hawaii Press.
- Hulstijn, J. and R. de Graaff  
1994 Under what conditions does explicit knowledge of a second language facilitate the acquisition of implicit knowledge? A research proposal. In *Consciousness in Second Language Learning*, J. Hulstijn and R. Schmidt (eds.), *AILA Review* 11: 97–109.
- Krashen, S.  
1982 *Principles and Practice in Second Language Acquisition*. Oxford: Pergamon Press.  
1985 *The Input Hypothesis: Issues and Implications*. London: Longman.  
1993 The effect of formal grammar teaching: Still peripheral. *TESOL Quarterly* 26: 409–411.
- Long, M. H.  
1991 Focus on Form: A design feature in language teaching methodology. In *Foreign language research in cross-cultural perspective*, K. de Bot, R. Ginsberg and C. Kramersch (eds.), 39–52. Amsterdam: John Benjamins.
- Long, M. H. and P. Robinson  
1998 Focus on form: Theory, research and practice. In *Focus on Form in Classroom Language Acquisition*, C. Doughty and J. Williams (eds.), 15–41. Cambridge: CUP.
- MacWhinney, B.  
1995 *The CHILDES Project. A Manual*. Hillsdale, NJ: Erlbaum.
- Montangero, J.  
1971 L'apprentissage de la voix passive. *Archives de psychologie* 41: 53–61.
- Norris, J. M. and L. Ortega  
2000 Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50 (3): 417–528
- Pica, T.  
1984 Methods of morpheme quantification: their effect on the interpretation of second language data. *Studies in Second Language Acquisition* 6: 19–78.  
1985 The selective impact of classroom instruction on second language acquisition. *Applied Linguistics* 6, (3): 214–222.

- Py, B.  
 1999 Enseignement, apprentissage et simplification de la langue. In *De la didactique des langues à la didactique du plurilinguisme, Hommage à Louise Dabène*, Billiez, J. (ed.). Grenoble: Presses de l'Université Stendhal.
- Richards, J. and T. Rogers,  
 1986 *Approaches and Methods in Language Teaching*. Cambridge: Cambridge University Press
- Robinson, P.  
 1995 Aptitude, awareness, and the fundamental similarity of implicit and explicit second language learning. In *Attention and awareness in foreign language learning*, R. Schmidt (ed.), 303–358. Honolulu: University of Hawaii Press  
 1996 Learning simple and complex second language rules under implicit, incidental, rule-search and instructed conditions. *Studies in Second Language Acquisition* 18: 27–68.  
 1997 Automaticity and generalizability of second language learning under implicit, incidental, enhance(d), and instruct(ed) conditions. *Studies in Second Language Acquisition* 19: 233–247
- Sabeau-Jouannet, E.  
 1975 Les premières acquisitions syntaxiques chez les enfants français unilingues. *La linguistique* 11: 105–122.
- Schmidt, R.  
 1990 The role of consciousness in second language learning. *Applied Linguistics* 11, (2) 129–58.  
 1993 Awareness and second language acquisition. *Annual Review of Applied Linguistics* 13: 206–226.  
 1995 Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In *Attention and awareness in foreign language learning*, R. Schmidt (ed.), 1–63. Honolulu: University of Hawaii Press.
- Sharwood-Smith, M.  
 1991 Speaking to many minds: on the relevance of different types of language information for the L2 learner. *Second Language Research* 7: 118–32.
- Sinclair, H. and E. Ferreiro  
 1970 Etude génétique de la compréhension, production et répétition des phrases au mode passif. *Archives de Psychologie* 40: 1–42.
- Skehan, P. and S. Foster-Cohen  
 1996 The influence of planning and task-type on second language performance. *Studies in Second Language Acquisition* 18: 299–323.
- Skehan, P. and P. Foster  
 1999 The influence of task structure and processing conditions on narrative retellings. *Language Learning* 49: 93–120.

- Spada, N. and P. Lightbown  
1993 Instruction and the development of questions in L2 classrooms. *Studies in Second Language Acquisition* 152: 205–224.
- Stern, H. H.  
1992 Issues and options in language teaching. Oxford: Oxford University Press.
- Van Patten, B. and T. Cadierno  
1993 Explicit instruction and input processing. *Studies in Second Language Acquisition* 15: 225–243.
- Westney, P.  
1994 Rules and pedagogical grammar. In *Perspectives on Pedagogical Grammar*, Odlin, T. (ed.). Cambridge: Cambridge University Press.
- White, J.  
1989 *Universal Grammar and Second Language Acquisition*, Amsterdam: John Benjamins.  
1990 Another look at the logical problem of foreign language learning: a reply to Bley-Vroman, *Linguistic Analysis* 20: 50–63  
1998 Getting the learners attention: a typographical input enhancement study, In *Focus on Form in Classroom Second Language Acquisition*, Doughty, C. and Williams (eds.), 85–113, Cambridge: Cambridge University Press.
- Williams, J. and J. Evans  
1998 What kind of focus and on which forms? In *Focus on Form in Classroom Second Language Acquisition*, Doughty, C. and J. Williams (eds.), 139–15. Cambridge: Cambridge University Press.





# **Focus on formS as a means of improving accurate oral production**

*Ron Sheen*

This chapter first examines fundamental aspects of the field of applied linguistics and second language acquisition (AL-SLA) in terms of the advocacy and fate of new approaches during the last half century. In doing so it endeavours to account for the apparent syndrome of failure to bring about identifiable improvement in classroom second-foreign language learning. Second, in the light of this syndrome, it submits to some scrutiny the development of the approach to grammar instruction known as focus on form (FoF) as opposed to a focus on formS (FoFS) and contends that its advocacy displays the same flaws of past failed advocacies. Third, it reports on an eight-month comparative study involving these two options the results of which demonstrate the greater effectiveness of a focus on formS in helping students to learn and use grammatical features. Fourth, it discusses implications of the findings of that study and of the previous discussion in terms of the future orientations of AL-SLA and the ensuing advocacies.

## **1. Introduction**

Since the 1950's, applied linguists have attempted to develop theories of second language learning based on the current theory of first language acquisition and have based their advocacies thereon. In more recent decades, such a theory of second language learning has been termed a theory of second language acquisition (SLA). Nevertheless, they have followed the same path as the previous applied linguists in basing reforms in language teaching on such theories. In order to avoid ambiguity, I will refer to the field in which these applied linguists function as applied linguistics-SLA (AL-SLA). In other words, this chapter is concerned with research and theories of classroom second and foreign language learning which have served as the basis for advocating a variety of instructional options in the last fifty years.

Three major reforms of the last half century have been of influence though the third one, task-based instruction (TBI) based on a focus on form (FoF), is still in-the-making (Brookes 1965; Howatt 1984; Dulay, Burt & Krashen 1982; Richards & Rodgers 1989; Long & Crookes 1992, Long 1991, Long 2000; Skehan 1996, 1998; Germain 1993). The first two reforms involve successively the introduction of the audio-lingual method (ALM) of the 50's and 60's based on a behaviourist approach, which was then replaced by strong communicative language teaching (SCLT) based on the principle of creative constructionism resulting in incidental learning of the 70's and 80's. Then, during the 80's, it became increasingly evident that SCLT was inadequate even though the underlying principle of creative constructionism was still considered valid. SCLT was, therefore modified to include an interactive component and a FoF thus leading to the creation of a new approach called task-based instruction (TBI) in which it is assumed that the learning process needs to occur during the performance of tasks relevant to the needs of the learners (see Long 1991; Long & Crookes 1992; Skehan 1996, 1998).

If we examine the nature of these three reform movements, we can identify certain common features. First, they were initiated by an assumption that the classroom results achieved by the then current approach could be improved upon. Second, that assumption was founded on the belief that that current approach was based on a flawed or less than adequate theory of SLA. Third, they were based on the assumed validity of the theory of SLA underlying the new approach. Fourth, applied linguists played a major role in such advocacies at conferences and in publications, thus persuading educational authorities of the legitimacy of the advocacy. Fifth, they all failed to submit their advocacies to successful long-term classroom trialling before advocating implementation, thus omitting to demonstrate the putative superiority of what they advocated.

## **2. The fate of such reforms**

It would be rewarding and encouraging had such reforms produced the promised improvement. This, unfortunately, is not the case, a conclusion justified by the absence of published empirical evidence demonstrating that any of these three approaches have lived up to their promise.

This is not surprising. Most reforms in education, in general, and foreign-second language learning, in particular, have proven to be failures (Adams and Chen 1981; Brumfitt 1981; Fullan 1982). In fact, Markee (1993: 231),

given the high risk of failure, argues that "...innovations should be resisted rather than promoted because their adoption may be more harmful than beneficial". Valette (1991: 325), indeed, argues, with supportive test scores, that the innovations of the previous twenty five years had resulted in the worsening of the proficiency standards of seniors graduating from college.

How then can one explain this history of failures? The answer appears to be reasonably self-evident and is largely implicit in the five common features mentioned above. In the first place, in spite of advocates' contending that what they propose will improve the learning of students, the field of applied linguistics has failed to address the issue of what we might reasonably expect learners to achieve given some well-defined situation and time-scale. If new approaches are claimed to improve learning, then these claims ought to be based on reliable argument supported by empirical evidence rather than merely on hypothetical prediction. Second, new approaches arise from the advocacy of a supposedly all-embracing new theory of SLA which, though possibly capturing some truths about foreign and second language learning, takes no account of the fact that many fluent speakers achieved their success without following the prescribed path of the new theory, thus throwing some doubt on its applicability to the classroom. Furthermore, though the advocacy may find some supportive findings in published research, other published research providing counter-evidence is frequently not referred to. More importantly, the new advocacy is not subjected to long-term trialling in normal classrooms demonstrating its efficacy. Further, the new advocacy is not called upon to prove itself in multiple rigorous comparative studies in order to demonstrate that what is proposed is indeed the most effective option.

### 2.1. Negative fallout from such reforms

Apart from the evident unfortunate effects of the failure of such reforms, there is another long term negative influence. Their implementation often results in the rejection of much that is associated with the approach to be replaced. This occurred when audiolingualism was introduced entailing the rejection of explicit grammar instruction and the use of explicit contrastive analysis and again when strong communicative language teaching (SCLT) replaced audiolingualism entailing the rejection of the principles of pattern practice, memorisation, the use of contrastive analysis of any sort and a concern with accuracy. This situation is commonly reflected in journals and conferences in the absence of treatment of such features. The third reform

based on a FoF differs somewhat in this respect and this, because it developed from SCLT and was thus evolutionary rather than revolutionary. Nevertheless, it has brought with it some of the unmotivated proscriptions developed by SCLT, particularly the stigmatising of a FoFS otherwise known as traditional instruction as in, for example, the grammar-translation method.

## 2.2. An explanation for the failures

Such a history of failure would in most fields result in serious soul-searching in order to discover the underlying causes and to find solutions. This, for the most part, has not occurred in applied linguistics. True, in some cases aspects of the past have been re-examined such that we might learn some lessons therefrom, examples being Diller (1975), Stern (1982), Howatt, (1984), Swan (1988), Spolsky (1989), Phillipson et al., (1991), Germain (1993), Musumeci (1998), Meara (1998). Further, mainline journals have recently devoted issues to reviews of the past (see Catford 1998 in *Language Learning* and Lantolf 2000 in *The Modern Language Journal*). A common theme of their reflections is the conclusion that a number of the strategies used in the past were wrongly rejected for, if used judiciously, they can make positive contributions to the learning process. This had already found some resonance in the return to respectability of the value of explicit contrastive input in language learning (see Gass & Selinker 1983; Kellerman & Sharwood Smith 1986; Odlin 1989). More recently, a special edition of *Studies in Second Language Acquisition* (2002-24) devoted to the effect on acquisition of frequent encounters with forms raises the possibility that yet one more rejected feature of language learning will regain its lost status. Though the edition is theoretically oriented, some of the conclusions drawn tend to indirectly support the use of repetition and practice in the classroom.

## 2.3. A failure to learn from the past

Sadly, this reflection on positive lessons to be drawn from the past implying criticism of developments in AL-SLA often does not seem to result in a return to past rejected practices. For example, though the work of Gass and Selinker and others mentioned above convincingly demonstrated the value of explicit contrastive awareness, it has had no effect on contemporary advocacy nor on the production of teaching materials which continue to be monolingual

making no mention whatsoever of the L1 of learners. Similarly, though the literature is replete with reports of the greater efficacy of teaching strategies based on explicit grammar instruction (see Smith 1970; Von Elek & Oskarsson 1973; Scott 1989; Palmer 1992; Kupferberg & Olshtain 1996; Sheen 1996; White, 2001; Erlam 2003), often the impression is created as if no such evidence existed or as if explicit grammar teaching, being “Neanderthal” (Long 1988: 136) in nature, was beyond the pale.

This failure to learn from the past is reflected in various state-of-the-art type papers, examples of which include Spada (1997), Robinson (1997), and Lightbown (2000). The first two are wholly concerned with the domain of a FoF, manifesting a somewhat narrow view of both the contemporary scene and the past. For example, in both articles, one finds little reflection on past practices such as the use of contrastive analysis and explicit grammar teaching. Nor does one find a critical scrutiny of the current advocacy of a FoF. Nor do these articles address what any such paper necessarily must address which is the degree to which a particular field has succeeded in respecting its implicit mandate. In the case of AL-SLA as defined in this chapter, it concerns the issue of the apparent failure of AL-SLA to bring about any demonstrated significant improvement in the efficacy of classroom language learning.

Some might argue that invoking this latter criterion is simplistic given that the overall aim of the field is to develop a theory of SLA in the long term. Certainly, this may well apply to those solely concerned with such an endeavour who make it crystal clear that their findings have no immediate relevance to the classroom. However, this is not the case for the applied linguistic research discussed and cited in this article. The principal purpose of this research is to justify the advocacy of what are considered to be optimal teaching and learning strategies. Furthermore, these researchers have not been loath to advocate the implementation of their findings in the classroom (see Long & Crookes 1992; Doughty & Williams 1998; Long & Robinson 1998; Lightbown 1998, 2002; Long 2000; Doughty 2001). As such implementation takes place in the real world, it has an effect on people’s lives. The issue, therefore, becomes an ethical one, thus making it essential that all applied linguists consider themselves accountable for what they advocate for classroom use. It is, then, incumbent on all involved to exercise great caution in advocating innovations and to, therefore, address the putatively “simplistic” question posed above. One essential element in answering this question concerns the failure to subject what is advocated to long-term trialling before proceeding to large-scale implementation. It is such – but post-hoc – trialling in addition to other means of evaluation that was used

in the assessment of the fate of reforms discussed above and which concluded that the implementation thereof had had a largely negative effect.

Applied linguists are, of course, aware of the pernicious effect of the implementation of doctrinaire approaches without prior long-term trialling. Larsen-Freeman & Long (1991: 290), for example, write “We must guard against overzealousness on the part of theorists or their devotees who feel that they have a monopoly on the truth”. Long & Robinson (1998) write with justifiable disparagement of “unproductive pendulum swings” and Long (2000: 179) ascribes some of the ills of the field “...to some drastic swings of the pendulum of fashion over the years”. However, the fundamental cause of such swings is the unquestioning acceptance of doctrinaire approaches putatively justifying the proscription of specific teaching practices – in other words, the very type of approach that is being advocated. For instance, the advocacy of a FoF (Long 1991; Long & Crookes 1992; Long & Robinson 1998; Long 2000) entails the proscription of a FoFS which Long (1988: 136) characterises as “Neanderthal”, whilst ignoring the fact that such an approach has facilitated countless learners’ achieving high levels of language proficiency (see, for example, Von Elek & Oskarsson 1973; Diller 1975; Strevens 1987; Chastain 1988; Spolsky 1989; Cook 1991; Stern 1992; Palmer 1992; Ur 1993; for a discussion of aspects of the effectiveness of such strategies). It should not, however, be assumed that all advocates of a FoF go that far. Doughty and Williams (1998b: 261), for example, in an indirect critique of Long’s position state, “...it is entirely possible to combine explicit and implicit FONF techniques, depending upon the particular acquisition circumstances. Such combinations should not be theoretically proscribed”.

In summary, then, in terms of the explanation of failures of past reforms, there would appear to be two major reasons. First, the reforms derive largely from theoretical argument based on the assumed legitimacy of a theory of SLA. Given this and, in addition, the fact that the new reform entails the proscription of strategies for which there is ample positive evidence in the literature, there is an obvious need to put that reform to some form of controlled long-term comparative trialling producing positive results before it is implemented. In the reforms discussed above, such trialling did not take place thus making failure a very probable outcome. This conclusion is supported by the post-hoc trialling which has taken place. In the case of ALM, there is Smith (1970) and Von Elek and Oskarsson (1973). In the case of SCLT, there is Scott (1989), Palmer (1992), Kupferberg and Olshtain (1996) and Sheen (1999), among others. In the case of FoF, it is as yet early days. However, so far no study comparing the effects of a FoF as opposed to a

FoFS has found in favour of the former whilst White (2001) and Sheen (2003), an interim summary of the study described in this chapter, provide convincing evidence of the greater effectiveness of a FoFS in terms of the learning and use of individual grammatical features.

### **3. The advocacy of a focus on form**

The FoF approach derives directly from the work of Long (1988, 1991, 2000) with the development of his interaction hypothesis, his re-establishment of the legitimacy of form-focused instruction and, more pertinently, his dichotomizing of such instruction into a focus on formS, which he proscribes, and a focus on form, which he advocates.

Put in broader terms and as previously discussed, the FoF approach originates indirectly from the less-than-impressive results achieved by the non-interventionist SCLT which refused any important formal role to grammar instruction, affording almost exclusionary priority to exposure to the language in the form of comprehensible input CI. Since then a number of applied linguists led by Long have proposed that CI is best experienced through interaction between learners and teachers and/or between learners and still contend that learners can acquire at least parts of the grammar thanks solely to this experience although Krashen, himself, has recently admitted that exposure solely to CI will result only in survival language skills bereft of redundant grammatical features (FLTEACH List debate with Krashen, 26 November 1999 and Krashen 1985).

This FoF approach to form-focussed instruction then became an integral part of TBI (Long & Crookes 1992). Since then, in spite of the fact that the interaction hypothesis has remained without empirical evidence in support of the claim that interaction alone results in acquisition of grammatical features and in spite of the fact that a FoF has nowhere been demonstrated to be the most effective option, a FoF and the task-based syllabus have been largely uncritically accepted by the field (yet see Burton 2002, Sheen 1993, 1994a and Swan forthcoming as notable exceptions to this).

#### **3.1. The failure to hold up a focus on form to critical scrutiny**

This is evident in various state-of-the-art papers, such as Spada (1997) and Robinson (1997) discussed above. It is also salient in Lightbown (2000). She claims that “[t]he influence of SLA research is now evident in textbooks



and teacher training programs and in proposals for curriculum design (see, e.g., Long and Crookes 1992)". While citing Long and Crookes (1992) in a positive light, she makes no mention of critical reviews of the Long & Crookes advocacy (see Sheen 1993, 1994a, and, indirectly, Skehan 1996; but see Long 1994 for a response thereto and Sheen 1994b for a rebuttal; see Lightbown 2002 for a recent explanation of her stance). Long & Crookes appeared in 1992 following the elaboration of the underlying principles of a FoF in Long (1988 and 1991; see also Long & Robinson 1998 and Long 2000 for recent formulations thereof). In spite of the fact that the long-term implementation of TBI has nowhere been demonstrated to be effective, its underlying principles are continually cited positively without demur (Doughty and Williams 1998a, 1998b; Doughty & Varela 1998; White 1998; Nassaji 1999; Lightbown 2000; Ellis, Basturkmen and Loewen 2001). The critical reviews thereof in the literature (Bruton 2002; Sheen, 1993, 1994a, 1994b, 2000, 2003) are not addressed. This has resulted in a climate in which researchers feel no obligation to submit a FoF to critical scrutiny and to justify their ignoring the claims to consideration of a FoFS (see Nassaji 1999, and Ellis, Basturmen, and Loewen 2001 as two illustrative examples among many). This is particularly surprising in the case of Ellis, Basturkmen, and Loewen (2001) for in Ellis (1994), he argues that "...it may be premature to reject a focus on formS approach" (p. 641). As since then there has been no published comparative study demonstrating that a FoF is more effective than a FoFS (in fact, the reverse is the case) and the findings of Norris & Ortega (2000) have been published (see below), there seems to be no justified reason for Ellis to modify that stance.

### 3.2. The implementation of reforms based on a focus on form

The influence of this uncritical acceptance has gone beyond the area of research studies, for it has now contributed to new reforms such as the ones recently enacted in Quebec, Canada. An examination of the sample materials used in this context shows them to be creative, interesting and engaging. There is, therefore, nothing to object to therein. What is objectionable, however, is the explicit proscription of a FoFS approach for it would appear that a FoF will be the only accepted means of teaching grammar. However, there is ample research demonstrating that grammar instruction may be implemented in CLT in a variety of differing explicit ways (Howatt 1984; Spada 1987; Palmer 1992; Carroll & Swain 1993; Celce-Murcia, Dornyei & Thurrell 1997). This broad range of different means of integrating grammar



instruction in CLT is also discussed in Doughty & Willams (1998a). Lyster (1998:186) further emphasises that the issue of the degree of explicitness of instruction remains "...the centre of much debate". This is, therefore, not the time to restrict the way in which form-focused instruction is implemented, particularly when there is no convincing research evidence to demonstrate that a FoF is the most effective way of doing so and when the only available published comparative research suggests a FoFS to be the more effective option (as will be shown shortly).

### 3.3. A critique of the advocacy of a focus on form

Given that much of the above has been devoted to criticism of the failure to subject to critical scrutiny the advocacy of a FoF, it is incumbent upon me to justify those criticisms by demonstrating the vulnerability of that advocacy to serious criticism. It is to this task which I now turn.

In order for the advocacy of a FoF to have any legitimacy, there needs to be convincing empirical evidence in support of its two central tenets: (a) that grammatical features may be acquired incidentally thanks ONLY to comprehensible communicative interaction and (b) that a FoF is the most effective means of making learners aware of the grammar where necessary. The last issue, (b), will be addressed in the following section, 3.4.

As to (a), Lightbown (2000: 439), for example, maintains that "[c]lassroom research has provided additional support for the conclusion that some features are acquired incidentally – without intentional effort or pedagogical guidance". Similarly, but more radically, Long (2000:179) states: "Given adequate opportunities, older children, adolescents, and adults can and do learn much of the grammar of a second language incidentally, while focusing on meaning or communication". Though, there is, indeed, little doubt that exposure to meaningful language may result in some degree of aural comprehension, whether it leads to accurate language production is another matter and a crucial one for as Spada and Lightbown (1993: 208) point out, it is "...spontaneous and free oral production tasks which provide a more accurate reflection of the learners' internal grammar". This being so, applied linguists who support a FoF would contribute to the debate were they to clarify whether they include both comprehension and production skills in their contention. Then, if they include the latter, it would be necessary to establish what would be required in terms of spontaneous production to support the interaction hypothesis. I would argue that to do so, minimally, one would have to demonstrate that accurate spontaneous production ulti-

mately manifests the existence of some native-like sub-system or other in that grammar. Thus, for example, the fact that learners are able to produce a question such as “Do you like soccer?” would not constitute proof of such acquisition for it may simply result from the process of memorizing chunks. What would be required, at least, would be the systematic spontaneous production of, say, the appropriate forms of some sub-system such as present interrogative forms of regular verbs. Or, of a less demanding nature, there would need to be evidence of the systematic spontaneous use of the indicative present tense forms to which the students have been exposed. If the students are able to do this, it would partially validate the interaction hypothesis. If, on the other hand, one were able to study the oral production of students who had been exposed only to CI and it transpired that they were only able to express appropriate meaning but in incorrect forms, one would have to conclude that the position is unfounded and that, in spite of extended exposure to English, learners are unable to acquire the sub-systems manifest therein.

If one examines the literature in search of empirical evidence in support of the interaction hypothesis in terms of the acquisition of grammatical features, one will find no evidence of the acquisition of a single grammatical feature demonstrated by spontaneous accurate oral production. In fact, contrary evidence has been implicitly provided by Lightbown et al. (2002), which gives a summary of the six-year New Brunswick comprehension-based ESL programme. That summary describes much of what the students produced as being error-ridden. More importantly, in all that has been written about this project, no hard evidence has been provided to support the reality of incidental grammar learning in terms of its bringing about accurate oral production. Thus, it is legitimate to consider that this central tenet of a task-based syllabus and a FoF is more a product of hypothesising rather than an empirically-supported fact of language learning.

As regards the issue as to whether learners can progress from some initial erroneous forms to correct forms. Lightbown has consistently argued that in acquiring grammatical features, learners pass through developmental sequences (see, for example, Lightbown 1998 and Lightbown & Spada 2002). However, supportive empirical evidence derived from her research on Quebec ESL learners has not been published. For example, in Lightbown and Spada (2002), it is claimed that learners progress from forms such as “What the dog are playing” to the production of correct forms. Yet, no empirical evidence derived from classroom incidental learning to support the claim is provided. This is understandable for there is, to my knowledge, at least, no evidence demonstrating that, on the basis of ONLY incidental

learning, such learners pass through increasingly correct forms until being able to produce native-like utterances.

In fact, there exists contrary evidence. In the study reported on below, the learners deprived of instruction on the formation of questions, spent three years producing this same “Wh- + no inversion” interrogative form. Further, in my own on-going cross-sectional research entailing several hundred interviews, I have followed students (subjected only to SCLT) who began as elementary students producing such erroneous forms and continued to do so through five years of high school and two years of college up to taking university admissions tests (Sheen 2005). This throws some doubt both on the reality of developmental sequences and on the legitimacy of the appeal to teachers to be patient while developmental sequences are allowed to run their course (Lightbown 2002: 534).

If we are to respect the available empirical evidence, we need to make clear to teachers that it does NOT demonstrate that students either deprived of pedagogical guidance or benefiting ONLY from corrective feedback do, in fact, go through such sequences but, rather, that they tend to fossilize forms such as the “Wh- + no inversion” interrogative forms. Further, on a more general level and given the lack of hard evidence of developmental sequences in the incidental learning of students in the above-mentioned six-year comprehension-based study in New Brunswick, there is an immediate need for all advocates of teaching strategies based on the putative validity of the incidental learning hypothesis, to demonstrate some accountability in reporting on their own research studies. These studies, do not, in fact, offer any evidence in support of that hypothesis. In other words, the absence of supportive evidence particularly in long-term studies needs to be discussed in the literature just as much as the putative presence thereof.

#### 3.4. Focus on form as opposed to a focus on formS

Before discussing the second central tenet, the assumed greater effectiveness of a FoF over a FoFS, the exact nature of these two options need to be clarified. Long (1991: 45–46) characterises FoF succinctly as an approach which “... overtly draws students’ attention to linguistic elements as they arise incidentally in lessons whose overriding focus is on meaning or communication”. This needs to be expanded and exemplified.

An underlying assumption of a FoF approach is that all classroom activity needs to be based on communicative tasks and that any treatment of grammar should arise from difficulties in communicating any desired meaning. Further,

that treatment preferably takes the form of quick corrective feedback allowing for minimal interruption in communicative activity. However, if on occasion there is a need for more extended grammar instruction, it is considered preferable to base it on grammar-problem-solving tasks and not on explicit language instruction. It is important to understand the implications of these features in terms of the overall syllabus. It means that a central planning feature is its task structure designed to answer the needs of the students. As the treatment of grammar depends on unplanned problems in communication arising during communicative activities, there is no grammar syllabus. Further, the only grammar to be dealt with is that which causes a problem of communication and not with a problem of form. If a student were to say, for example, "Your wife, where she go tomorrow?", this would, in fact, not justify interrupting the on-going activity because the meaning is unambiguous. On the other hand, if a francophone student were to say, "She gave me his book.", it is feasible that the student actually meant to say "her book", the mistake arising from the difference between possessive determiners in English and French. As this error causes a breakdown in meaning, the teacher would be justified in devoting time to bringing the students to an understanding of this difference and this, by means of problem-solving tasks (see White, 1998, for an example of a FoF using such tasks and for her conclusion that a more explicit teaching of the contrastive problems involved would have been preferable.)

In the FoF literature, a FoFS tends to be characterised negatively as a traditional forms-in-isolation type of grammar teaching. This, however, constitutes a misrepresentation if one perceives of a FoFS as described in the work of Chastain (1970, 1988), Ur (1993), Celce-Murcia, Dornyei & Thurrell (1997), Dekeyser (1998), Sheen (1996, 2003) and Johnson (2001). Their perception of a FoFS shares with that of a FoF the assumption that communicative activity is the underlying priority of the classroom. However, it assumes that given the great difficulty of learning the grammar and vocabulary of a foreign language, these can be learned effectively neither incidentally as a by-product of communicative activity nor simply by means of problem-solving activities. It proposes rather a skills-learning approach which comprises three stages (see DeKeyser 1998 and Johnson 2001 for a detailed description): (1) providing understanding of the grammar by a variety of means, including explanation in the L1, pointing out differences between the L1 and the L2, and aural comprehension activities intended to focus students' attention on the forms being used; (2) both written and oral exercises entailing using the grammar in both non-communicative and communicative activities; (3) providing frequent opportunities for communicative use of the

grammar to promote automatic and accurate use. Though it assumes that this is the most effective means of teaching grammar, it does not preclude any of the techniques used in a FoF as it adopts an eclectic approach to classroom activities. Such eclecticism is also evident to some degree in advocacies of a FoF. Thus, Lightbown (1998) and Doughty and Williams (1998b) agree that a FoFS may have its use with certain complicated grammatical features. Nevertheless, they all give the strong impression that a FoF is the default option.

### 3.5. The empirical evidence in support of a focus on form

We now return to the critique of the advocacy of a FoF. When one examines the arguments used to support this proposal (Long & Crookes 1992; Long & Robinson 1998; Long 2000) one notes two features. First, the underlying argument is largely theoretical, being based on the assumed validity of the interaction hypothesis and the related theory of SLA, and that the empirical evidence offered in support is weak, at best, and flawed, at worst. Doughty & Varela (1998) exemplify the first, in stating: “Therefore, in our view, a quintessential element of the theoretical construct of focus on form is its dual requirement that the focus **MUST** (my emphasis) occur in conjunction with – but **MUST** (my emphasis) not interrupt – communicative interaction” (p. 114).

In order to provide empirical support for such a position, there would ideally be available the findings from comparative studies demonstrating Long’s option of FoF as being more effective than, say, something related to a FoFS. However, as Lightbown (2000) reveals, there is virtually no comparative evidence to support either. Furthermore, while there is, at best, only limited empirical evidence of the successful implementation of the FoF approach in real classrooms, there is an abundant amount of empirically-supported argument in the literature demonstrating the positive degree to which a FoFS may be instrumental in enabling learners to reach accurate fluency in FL’s (see, for example, Von Elek & Oskarsson 1973; Palmer 1992; Kupferberg & Olshtain 1996; Sheen 1996, 2003).

Given this situation, one might expect some prudence in the literature in the use of either terms, FoF and FoFS. This is not the case. A number of the contributors to Doughty and Williams (1998a), including the editors themselves, and Lightbown (2000) use the term FoF sometimes in Long’s sense, sometimes in a broader sense. At the same time, the impression is created as if the argument in favour of the Long option has already been demonstrated

to be valid and that a FoFS has already been shown to not be the most effective option (see, however, the summarising chapter of Doughty & Williams 1998b, for a more balanced view).

However, as argued earlier, the evidence offered in support of a FoF is unconvincing. What is consistently cited as empirical support are various papers by Lightbown and Spada who, much to their credit, have been instrumental in publishing findings on Quebec classroom-based research. On the basis of their own work, Lightbown, for example, states: "Research in intensive ESL classes with young francophone learners has shown that teachers who focus learners' attention on specific language features during interactive, communicative activities of the class are more effective than those who never FoF or who do so only in isolated grammar lessons (Lightbown, 1991; Lightbown & Spada, 1990; Spada & Lightbown, 1993). These effective teachers tend to provide FoF on the fly, without causing the interaction to be interrupted or learners to be discouraged" (Lightbown 1998: 192).

It is revealing to submit this citation and the research to which it refers to some analysis because it is so frequently quoted in support of a FoF. It is problematic in a number of ways. First, the research referred to is not ONLY on ESL intensive classes but also on regular classes of two hours a week of audiolingual sessions (see Lightbown, 1991: 203). Thus, two entirely different forms of instruction from two different studies are being compared, which detracts from the validity of the conclusion reached. Furthermore, it is stated that because the drills used in the audiolingual lessons devoted to teach students to use "there's" did not prevent their over-using "have" once they learned it, they were not effective. This finding is then used to conclude that separate grammar lessons are not as effective as other instructional options. This conclusion is premature. If one wishes to evaluate the efficacy of separate grammar lessons, one does not take audiolingual drilling as examples of such lessons. The method comparison research of the 60's and 70's overall showed audiolingual instruction to be less effective than cognitive-code-learning type instruction (Smith 1970, Von Elek and Oskarsson 1973), a form of the skills-learning option (DeKeyser 1998). One needs, therefore, to take a more effective option to represent separate grammar instruction before dismissing it out of hand as being not sufficiently effective. Then, in order to carry out a rigorous comparison, apart from the obvious one of comparing like-groups in like-situations (which is not done in the research cited), one would take account of the difficulties caused by the L1 interference of other forms. Then one would compare the learning outcomes of such a treatment with that of another group benefiting from another



instructional treatment. In the case in point, Lightbown compared the learning outcomes of a regular audiolingual group of some years previously with those of an intensive ESL class without indicating how much time each group spent on the grammatical point – an immediate disqualifier, particularly given the contrast between the appreciably briefer class time of the audiolingual class compared with the ESL intensive class. Furthermore, the “instruction” of the ESL class in question was ostensibly limited to instruction providing corrective feedback. This account is based on the anecdotal evidence of a single teacher whose providing of that corrective feedback was not actually observed by Lightbown or her collaborators. We can therefore only assume, as Lightbown does, that the “drumming into their little heads” (1991: 207) was carried out “...on the fly, without causing the interaction to be interrupted or learners to be discouraged”. However, not having observed the classes in question, Lightbown’s conclusion seems unjustified, or at least premature. Indeed, the editors, Doughty and Williams, in their final chapter, characterize this supposed intervention “on the fly” somewhat differently. They state: “Her technique was to correct learner errors and remind students of a metalinguistic lesson (ie separate grammar lessons) that she had apparently provided at an earlier class session” (Doughty and Williams 1998b: 207). Thus, Lightbown, whilst being somewhat disparaging of separate grammar lessons, is, at the same time, citing as an example to emulate, the instruction of a teacher who based her corrective feedback on such separate lessons previously given.

Yet, this particular “finding” is continually cited without qualification as support for a FoF (see, for example, chapters in Doughty and Williams 1998a). Other studies are also often cited in support of a FoF. Lightbown (2000) cites, in addition to the ones above just discussed, Doughty & Varela (1998), Harley (1989), and Lyster (1994), whose “...results have provided strong support for the inclusion of focus on form (not ‘formS’ my emphasis and addition) in the CLT classroom” (Lightbown 2000: 445). However, these studies justify no such conclusion if “focus on form” is being contrasted with “focus on formS” as it apparently is. The students in Doughty and Varela (1998) received FoFS in their language arts classes which they almost certainly used in responding to corrective feedback in the content classroom. Harley (1989) entails a FoFS for it entails a variety of activities devoted to bringing students’ attention to tense and aspectual forms but not to meaning problems raised in interactive communication. Lyster (1994) focuses on the issue of the efficacy of interaction involving form but does not address the issue of a comparison between the effectiveness of a FoF and a FoFS.

Let us be clear as to what is at stake here. If we accept, as I think many applied linguists do, that students need to be provided with grammatical knowledge in the service of the need to communicate with some degree of accuracy, we need to find the most effective means of providing that knowledge. The choice lies between accepting some exponent of the restrictive FoF as opposed to some exponent of a FoFS as in a skills-learning approach which, whilst accepting that the ultimate purpose is to communicate, proposes that the instruction may be separated from communicative activity with the intention of enabling students to master the new grammar in terms of both meaning and formal manipulation, before integrating it into communicative activities (see, for example, DeKeyser 1998, for such proposals).

Given the lack of comparative studies involving a FoF and a FoFS and the continuing controversy concerning their relative merits, two applied linguists, Norris & Ortega (2000), have produced a meta-analysis of many of the relevant studies. They concluded that a "...a focus on form and focus on formS are equally effective". However, though the study is highly sophisticated, the argumentation putatively justifying this conclusion, reveals some weakness. Though these two authors use Long's term, "focus on form", they do not use his exact criteria to differentiate a "focus on form" from a "focus on formS". For example, they classify VanPatten and Sanz (1995) as an example of a "focus on form" though an essential feature of their approach entails an explanation of discrete items of grammar unconnected to a need created by communicative activity, thus disqualifying it as "focus on form" in Long's terms. Moreover, Norris and Ortega utilize criteria which entail the exclusion of all pre-1980 comparative studies and a number of puzzling omissions among those published after that date (Kupferberg & Olshtein 1996; Palmer 1992; Sheen 1996, all of which found a FoFS to be the most effect instructional option). Their conclusions can, therefore, not be considered as an endorsement of Long's FoF. More specifically, their conclusion (p. 501) that "...a focus on form and focus on formS are equally effective" should be treated with scepticism, because, first, had Norris and Ortega used Long's criteria to decide what does and what does not constitute a FoF, and second, had they included all relevant studies, this particular finding would have favoured a FoFS. In spite of this, leading applied linguists have cited the findings, paying no heed to the factors mentioned above. Long (2001:189), for example, incorrectly maintains that Norris & Ortega's findings show that a FoF is superior to a FoFS.<sup>1</sup>

The various weaknesses in the research discussed above, the apparent bias in favour of a FoF in Norris & Ortega (2000), the less-than accurate representation of Norris & Ortega's findings in the recent literature and,



most importantly, the failure by advocates of a FoF approach to include an exponent of a FoFS in their studies make it essential to undertake comparative studies entailing the evaluation of the effects on learning of these two options. This is, in fact, exactly what Long himself proposed when he introduced the distinction between FoF and FoFS: “True experiments are needed which compare rate of learning and ultimate level of attainment after one of three programs: focus on forms, focus on form, and focus on communication... Research has yet to be conducted comparing the unique program types” (Long 1991: 47–48).

This brings us to a crucial part of this chapter: a description and discussion of such an experiment.

#### **4. A comparative study of the effectiveness of a focus on formS as opposed to a focus on form**

However, before coming to the study itself, various issues need to be addressed. Ideally, such a study would be long-term and have, at least, two groups equal in characteristics such as aptitude and motivation and having no previous experience in the second-foreign language being learned. Such conditions are not easy to find. Further difficulty arises when the study entails exposing the two groups to different teaching conditions. Schools now have pre-determined syllabi which they are ethically obliged to respect. It is, therefore, impossible for a school to accept a long-term study which does not respect the set syllabus with one group of students without the permission of all the parties concerned.

As it is often difficult to find groups of true beginners, many researchers are obliged to use subjects who have already learned the language for several years. No matter how many pre-tests one conducts, one can never be certain that previous learning and experience have not influenced the ultimate findings. This problem particularly applies to non-beginners who have been ostensibly subjected to SCLT and who have, therefore, not benefited from any grammatical instruction. Often, teachers in such programmes, frustrated by the lack of understanding of their students, resort to explicit grammatical instruction to solve their problems. This is the situation which exists in intensive ESL classes in Quebec elementary schools.

A further problem is caused by the fundamentally different nature of a FoF as opposed to a FoFS. The former does not systematically target specific grammatical features as does a FoFS. It is, therefore, unjustified to evaluate the effects of a FoF during a short period of no more than a few weeks.

Given these thorny problems, I was very fortunate, in 2001, to come across an elementary school in Quebec in which the sole English teacher had implemented the SCLT programme as intended and had thus insisted on a policy of English only and had not used grammatical instruction other than providing occasional corrective feedback. All his classes were based on the principle of engaging the students in enjoyable task or game activities in which they were only allowed to use English. He, himself, addressed them in English without making any attempt to modify his delivery to help the students understand. In fact, in the many classes I observed, I found that he spoke so quickly that I was pretty sure that even the best students could only get the merest gist of what was said and not the full meaning. Moreover, a good number of students remained passive and did not venture speaking in English. Furthermore, what English I heard from the students was, for the most part, error-ridden. However, the overwhelming impression was of two groups of young students with great affection for their teacher enjoying all the activities he devoted so much energy to creating. It was also impressive to have some of his students approach me to ask questions in English, albeit in incorrect forms.

Discussion with the teacher proved productive as he agreed to collaborate with me on both evaluating his two sixth grade classes and in carrying out an experiment with those same students. It is to the details of this study which we now turn.

#### 4.1. Purpose

The purpose of this study is threefold:

- a) To evaluate in an initial oral interview, the oral production of a whole range of interrogatives and utterances containing frequency adverbs with a view to testing the validity of the hypothesis that learners are able to acquire much of the grammar of the language they are learning thanks only to exposure to comprehensible input in interactive situations.
- b) To compare the effects on the oral production of interrogative forms and utterances containing frequency adverbs of two different treatments: a FoFS, and SCLT combined with a FoF.
- c) To evaluate the apparent effects on the underlying grammatical competence of the learners of the two treatments.

#### 4.2. Subjects

Two groups of 6<sup>th</sup> grade elementary students aged between 11 and 12 participated in the study. They were of middle-class background and, being in an enriched programme, enjoyed the strong support of their parents. One of the classes contained 30 students and was designated the control group (CG) for it continued basically as it had done for the previous two years with the addition of a FoF in the case of the grammar taught to the experimental group. The other class contained 18 students and was designated the experimental group (EG) and was taught by a FoFS. All students in the study had already had the same teacher for the previous two years.

#### 4.3. Timetable

Both groups had four contact hours each week. The researcher (myself) had 75 minutes class time with the EG with the rest of the time being given by the class teacher. The class teacher was responsible for all the class time with the CG except for the lessons when I conducted oral interviews with that group. The period of the actual teaching lasted from the beginning of October 2001 to the end of April 2002 with a break during the Christmas period. In the period up to Christmas, the teaching was concerned only with interrogative forms. In the subsequent period, the teaching of adverb placement was introduced while the revision and teaching of interrogative forms continued. Teaching ended at the end of April; however, the study continued until June when the last posttest was taken.

#### 4.4. Testing materials

Two types of instruments were used to measure the students' mastery of the target forms: an aural written comprehension test and an oral interview.

The aural written comprehension test consisted of a text of a hundred words and a list of 12 questions requiring information answers (see Appendix 1).

The first set of oral interviews differed from the subsequent ones. Their purpose was two-fold. The first purpose was to encourage the students to produce as much language as possible by asking them questions about all aspects of their daily lives, their families and their past life. The second purpose

was to get the students to ask ten questions which they had been asked in the first part of the interview. They are as follows:

1. Where were you born?
2. What did you do last night?
3. What are you going to do tonight?
4. Where does your father work?
5. How long has he lived here?
6. Do you like winter?
7. Does your mother prefer summer?
8. What is she going to do tonight?
9. What does your father do every Saturday?
10. What did he do last night?

The students had to attempt to produce these questions in order to discover the information they were asked to obtain in written instructions in French.

The content of subsequent interviews was limited to the interviewer's asking the same questions that the students were then instructed to ask of the interviewer. Those questions were of similar types to those used in the first interviews but for the most part were not the same.

In the post-Christmas period interviews, the students were instructed in both English and French to produce utterances containing the following adverbs: *sometimes, often, frequently, never, usually, rarely, seldom*. Such an instruction would be, for example, "Dites-moi que vous allez souvent au cinéma" ('Tell me that you go to the cinema often'). Or "Tell me that your father seldom goes to the movies".

#### 4.5. Teaching materials used for the EG

The teaching materials consisted of two sets, one for the teaching of the interrogative forms and the other for the teaching of correct adverb placement.

Set 1 consisted of four different types of activities:

- a) The following chart which was built up for the different interrogative forms as they were explained and the patterns understood.

Where	were	you	born
Where	was	person	born.
What	do	you	verb
Where	does	person	verb
Where	did	person	verb
What	is	person	going to + verb
Where	are	persons	going to + verb
Where	are	persons	going to + verb
How long	have	you	verb+ed
How long	has	person	verb+ed
Since when	have	you	verb+ed
Since when	has	person	verb+ed

- b) Oral pair work material of the following type in both English and French:

1. Ask me where I live. ‡ Where do you live?
2. Ask me where my friend lives. ‡ Where does your friend live?
3. Demandez-moi si j'aime le soccer. ('Ask me if I like soccer') ‡ Do you like soccer?
4. Demandez-moi ce que je vais faire ce soir. ('Ask me what I am going to do tonight') ‡ What are you going to do tonight?

- c) Regular homework containing 20 items of the following type:

Demandez-moi ce que vont faire mes amis ce soir. ('Ask me what my friends are going to do tonight'), requiring the answer "What are your friends going to do tonight?"

- d) Task work requiring the students to obtain prescribed information of their fellow students by circulating amongst them and asking questions.

Set 2 consisted of three types of activities:

- a) A chart containing sentences such as “I usually play soccer on Saturday” and “My mother never goes to the movies”.
- b) Oral pair work material containing items such as:  
Dites moi que vous allez rarement au cinéma. (‘Tell me that you rarely go to the movies’).  
Tell me your father often plays baseball.
- c) Homework containing 20 items similar to the French items in (b) above.

#### 4.6. General procedures

In the week before the actual study began, students in both classes took the aural comprehension test and underwent an oral interview.

Subsequently, the researcher taught the EG for a single lesson each week. Each lesson was devoted to three types of activity:

- a) Explanation given in French of the difference between French and English of the structure of questions and the building up of the chart given in the teaching materials. Using “Ask me ...” and “Demandez-moi...” type questions.
- b) Pair work using the same type of questions. This pair work functioned in the following way. Each student received a list of instructions of the type: “Ask me if my mother likes movies” and-or “Demandez-moi ce que j’ai fait hier soir” (‘Ask me what I did yesterday evening’). Next to each instruction was given the question form that the partner should produce thus permitting the other student to provide guidance and correction.
- c) Task work entailing students asking questions in order to obtain pre-scribed information about their classmates.

As the weeks passed, the time devoted to explanation reduced as the time devoted to activities increased. At the end of the pre-Christmas period, both classes underwent oral interviews in mid-December as their first posttest. In the first week of the post-Christmas period at the end of January, students in both classes underwent an interview to test retention of interrogative forms and to test their ability to produce utterances with correct adverb

placement. These tests therefore functioned respectively as a second post-test on interrogative forms and a pre-test on adverb placement.

As adverb placement does not lend itself to “Ask me...” type questions, in this case items such as “Tell me that you often go to the movies” were resorted to. However, as the use of English in such items reveals to the students the correct adverb placement, these items were asked in French as in, for example, “Dites-moi que vous allez souvent au cinéma” (‘Tell me that you often go to the movies’).

The subsequent weekly lessons now entailed continuing work on interrogative forms in pair and task work with the addition of further explanatory and pair work on adverb placement.

In mid-April, students in both classes underwent further posttests in the form of oral interviews entailing producing interrogative forms and utterances containing adverbs. This was the end of the period of the study. The experimental group then returned to the same instruction of the CG and thus had no further work directly intended to promote practice of interrogative forms and correct adverb placement.

At this point, both groups took a written grammaticality judgment test containing 18 items containing interrogative forms, 6 items containing adverbs and six declarative sentences as distractors (see Appendix 2). During the same week, both groups took once again the aural written comprehension test which both groups had taken at the beginning of the study.

At the end of June, both groups underwent oral interviews as their final post-test which followed the same procedure as the others.

#### 4.7. Scoring procedures

In the case of the interrogative forms, students were awarded one point if they produced the correct word order with the correct choice of auxiliary. If they failed to produce this correct word order, they received a zero score. Thus, an interrogative form such as “Where are you going tomorrow?” would receive one point whereas forms such “Where you go tomorrow?” and “Where you going tomorrow?” would receive zero points. In the case of adverbs, correct placement was awarded one point and incorrect placement a zero score.

In the grammaticality judgments, the identification of a correct form received one point. The identification of an incorrect form received one point only in cases where the students made appropriate corrections. Thus, in the case of item 13, “How many times your sister live here?”, identification as

incorrect followed by the correction, "Since when your sister lives here?" would not receive a point and this, because the student has failed to demonstrate knowledge of the correct question form in such constructions. On the other hand, a student who produced a correction such as "How long has your sister live?" did receive a point in spite of the error in "live" and this, because she demonstrated correct knowledge of the required syntactic construction.

As for the aural written comprehension test, the same evaluation procedure used at the beginning of the study was applied one again. That is: one point for a correct response, zero for an incorrect response.

## **5. Summary of results**

The initial pre-tests entailing the aural written comprehension test (see Appendix 1) and oral interviews in October 2001 indicated first, that both classes were of a similar standard with the CG being marginally stronger as the teacher had intimated before the study began, and, second, that whilst the students had good aural comprehension they were extremely weak in terms of accurate oral production.

The subsequent posttests for the CG showed no significant change in accurate oral production. The same tests for the EG showed dramatic and significant improvement in oral production at the end of the two instructional periods in December 2001 and April 2002. On the other hand, the post-tests following the non-instructional periods in January and June 2002, showed a reduction in accuracy (see results below). However, more importantly, the results still showed significant improvement compared with the pretest.

As to the aural comprehension tests taken in April 2002, the results were very similar to those of the pre-test in October 2001, the only difference being that the EG achieved slightly better results than the CG, the reverse of the results obtained in October 2001.

The results of the grammaticality judgment tests taken in April 2002, revealed a significant difference in favour of the EG. It is an unfortunate oversight on my part that such a test was not part of the pre-test. However, given that the other pretests taken in October 2001 revealed no significant difference in aural comprehension and oral production, it is safe to assume that had a grammaticality judgment test been taken at that time, it would have revealed no significant difference between the two groups and that the instructional treatment given to the EG explains the significant difference between the EG and the CG.



## Statistical analysis:

The data derived from the oral tests on question forms and adverbs were analysed separately by an analysis of variance following a two-dimensional (groups x moments) ANOVA with repeated measures on the second dimension. The data consisted of binomial variables (number of correct responses for a number of 10 items). Though not normally distributed, such data has proven to be both suitable and robust for normal-distribution procedures such as t-tests and analyses of variance. On the other hand, the suitability in the present case might be questioned given the consistently low performance of the students in the CG where the individual results remain in the 10% to 20% region. This objection may, however, be neutralised by the sustained difference through multiple posttests between the results of the CG and EG which indeed guarantees the reported significance.

The raw scores derived from the oral interviews as pre-tests and post-tests for question forms and adverbs are given below.

*Table 1.* Mean raw scores and standard deviations for question forms

	<b>Focus on form (N=30)</b>		<b>Focus on formS (N=18)</b>	
	M	SD	M	SD
Pretest	1.000	0.377	0.833	0.618
Posttest 1	1.033	0.414	7.389	1.852
Posttest 2	0.967	0.718	4.944	1.798
Posttest 3	1.133	0.732	7.556	1.886
Posttest 4	1.267	0.784	5.167	1.724

*Table 2.* Mean raw scores and standard deviations for adverbs

	<b>Focus on form (N=30)</b>		<b>Focus on formS (N=18)</b>	
	M	SD	M	SD
Pretest	0.100	0.305	0.056	0.236
Posttest 1	0.100	0.305	6.833	1.383
Posttest 2	0.100	0.305	4.278	1.606

The ANOVA with repeated measures carried out on the results of the oral interviews revealed a global advantage for the EG on question forms and adverbs (Questions:  $F=182.10$ ;  $DF=1$  and  $46$ ;  $p<0.01$  . Adverbs:  $F=357.78$ ;  $DF=1$  and  $46$ ;  $p<0.01$ ). It also revealed significant progress from pretest through all posttests (Questions:  $F= 301.45$ ;  $DF=4$  and  $184$ ;  $p<0.01$  . Adverbs:  $F= 429.67$ ;  $DF=2$  and  $92$ ;  $p<0.01$ ).

Also revealed were significant differences between the two groups on all posttests:

Questions: Pretest:  $t = -0.246$ ,  $df=46$ , non-significant  
 Posttest 1:  $t = 9.361$ ,  $df=46$ ,  $p<0.01$   
 Posttest 2:  $t = 5.857$ ;  $df=46$ ,  $p<0.01$   
 Posttest 3:  $t = 9.459$ ,  $df=46$ ,  $p<0.01$   
 Posttest 4:  $t = 5.744$ .  $df=46$ ,  $p<0.01$

Adverbs<sup>2</sup>: Pretest:  $t = -0.400$ ,  $df=46$ , non-significant  
 Posttest 1:  $t = 61.209$ .  $df=46$ ,  $p<0.01$   
 Posttest 2:  $t = 37.982$ ,  $df=46$ ,  $p<0.01$

Analysis of the Grammaticality Judgments task (April 2002) yielded the following results:

	N	Mean	Std. Dev	Std. Err
Control	30	9.000	3.126	.602
Experimental	18	17.000	3.202	.776

A post-hoc Scheffé test with significance level 5% provided the following:

	Mean diff.	Critical diff.	P-value
CG-EG	8.189	2.018	<.05

These results reveal a significant difference between the CG and EG in favour of the EG ( $t = 8.189$ ;  $df = 46$ ;  $p < 0.01$ ).

## **6. Discussion**

Before embarking on a discussion of these results, I repeat here the aims of this study:

- a) To evaluate in an initial oral interview, the oral production of a whole range of questions and utterances containing frequency adverbs with a view to testing the validity of the hypothesis that learners are able to acquire much of the grammar of the language they are learning thanks only to exposure to CI in interactive situations.
- b) To compare the effects on the oral production of questions forms and adverb placement of two different treatments: a FoFS and SCLT combined with a FoF.
- c) To evaluate the apparent effects on the underlying grammatical competence of the two treatments.

In terms then of the three purposes of the study and the results described above, the following conclusions are justified.

- a) The pretests on both question forms and frequency adverb placement revealed that none of the students in either group had acquired either feature and this, in spite of being exposed to CI during the two previous years.
- b) The posttests of the EG both for question forms and adverb placement revealed a dramatic improvement and, therefore, a statistically significant result (but see comments below).
- c) As far as grammatical competence is concerned as revealed in a grammaticality judgment test, the students in the EG demonstrated a much greater awareness of the relevant underlying grammar than did the students of the CG, thus producing a statistically significant result. Of particular interest here are the differences between the responses made by the CG as opposed to those of the EG in the case of third person interrogatives. In the case of number 27 (see Appendix 2), “How long have your brothers lived here?”, 11 out of 30 of the CG marked it correct as opposed to 14 out of 18 of the EG. Of further interest is the type of corrections made by those who marked it as being incorrect. Of the four students of the EG who did so, three retained the word order they had been taught during the study, replacing ‘have’ with “was” or “are” whilst one returned to the word order they had developed before the study

began. That is, “How long your brothers lived here?”. On the other hand, of the nine students of the CG who marked this number incorrect and also offered a correction, eight repeated “How long your brothers lived here?” with only one student retaining the correct word order, replacing “have” with “did”. This is of some importance for it reveals that when students are continually allowed to use incorrect forms without benefiting from explicit instruction as to the correct form, they will come to consider the oft-repeated incorrect form as being correct.

Before discussing the significance of the results obtained, a number of points need to be made in terms of the comparative nature of the study. That is, in terms of the acquisition of question forms by the CG the study did not allow for the direct emphasis on such forms by the teacher using a FoF approach. All that he was requested to do during his teaching was to provide corrective feedback should his students make errors in question forms or adverb placement. However, though his students had the need to use these forms on occasion during the task work they carried out, there was no provision for them to produce the specific forms used in the study. Thus, to take but one example, there is little doubt that the students in the CG had no need to produce a question such as “How long has your father lived here?” during the period of the study. Similarly in the case of adverb placement, no attempt was made to have the students in the CG use the necessary structures though they did so on occasion. These two affirmations are confirmed by the teacher of the CG.

Given this, there would be no justification to state that the findings of this one study demonstrate that a FoFS approach is more effective than one based on a FoF. However, taking the literature as a whole and studies aimed at evaluating the effect of a FoF, there is every reason to take seriously such a comparative statement. If one takes a number of studies attempting to put the effects of a FoF to the test, one finds there is little therein to inspire confidence in the positive effects of such an approach. If one examines, for example, a number of studies which attempted to evaluate the effect of a FoF on various grammatical features (White 1991; Spada and Lightbown 1993; Doughty and Varela 1998; White 1998), nowhere does one find anything akin to the marked improvement achieved by the FoFS in this study. Of further significance is White (2001). In her previous study, White (1998) had compared the effects of multiple exponents of a FoF concluding that given the poor results achieved, it would have been more effective to have used a FoFS approach exploiting contrastive input. In fact, she did this in her second study (White, 2001) comparing the effects of a FoFS with a FoF,

finding that the former produced markedly more significant positive results than did the latter. In fact, her comparisons between her CG and EG reflected the same dramatic differences found in this study

Further, if one examines the findings of other studies which have compared the effects of a FoFS with various forms of indirect instruction akin to a FoF (Palmer 1992; Kupferberg and Olshtain 1996; Sheen 1996), one discovers that a FoFS consistently produced significantly superior results to those achieved by the other approaches used. In fact, if one goes back to earlier comparative studies (see Von Elek and Oskarsson 1973 for a survey) one finds once again that approaches exploiting a FoFS consistently produce better results than other approaches.

In summary, then, it is fair to conclude that the application of instruction based on a FoFS, providing opportunity for frequent oral practice, enabled the students to demonstrate dramatic improvement in oral production entailing question forms and correct adverb placement in controlled oral practice.

In spite then of the caveats discussed above, these are results of some significance given the marked tendency of advocates of a FoF to dismiss the value of a FoFS. However, I am not suggesting here that the oral interviews permitted a reliable evaluation of the spontaneous use of question forms and adverb placement. In fact, in subsequent task-work requiring students to ask questions under time-pressure of their fellow students, appreciably fewer correctly-formed questions were asked. Nevertheless, what is of significance is that a good proportion of the students who in their pre-test in October 2001 produced no correct questions, in such task work, the EG students produced under time pressure some quite difficult questions correctly. What I am suggesting, therefore, is that a FoFS approach allowed the students to understand the underlying grammar and then use it in producing accurate language orally. Further, it enabled some of them to produce those forms in near-spontaneous conditions. However, at the same time, it needs to be made clear that the weaker students had great difficulty in such situations, manifested in the occasional combining of two auxiliary verbs in the same question such as "What do you will do...?" thus providing substance to the both persuasive and rather obvious argument that different types of students fare best with different types of teaching and learning strategies.

However, whatever the success-rate of the students, as all teachers know, the passage from declarative knowledge and use under controlled-exercise conditions to truly spontaneous use in natural communicative situations constitutes a major obstacle. Following a skills-learning approach, this

step entails providing the students with frequent opportunities to practise these forms in communicative situations in order to negotiate this major hurdle and achieve automaticity.<sup>3</sup>

To return to the results of the present study, this on-going research constitutes a qualified endorsement of a FoFS approach. It helped the students in the EG to make solid progress in the two targeted grammatical areas whilst the CG, benefiting only from a great deal of interactive comprehensible input combined with FoF in the shape of corrective feedback in the target areas, continued producing largely incorrect forms thus allowing fossilization to continue to develop. In fact, they achieved no recognisable progress in the target areas during the eight months of the study. However, to be fair to the CG, the testing did not involve global ability which is the target of the task-based approach used by the class teacher. But, then again, as the initial pre-tests demonstrated that that teaching approach had not, after over two years of learning, enabled the students to orally produce accurate language, this effect of a FoFS approach is of some significance. All this said, however, given the tentativeness of the findings of any single study in terms of external validity, there needs to be multiple replication of studies comparing the effectiveness of a FoFS and a FoF. It can be seen that what these findings and those of other studies involving a FoFS do is to demonstrate the justification of the essential inclusion of a FoFS in future comparative studies.

## **7. Summary and conclusion**

In this chapter, it has been noted that the major reforms in second and foreign language teaching of the last fifty years or so have largely failed to live up to their promise. It has contended that a major factor in this history of failure has been the tendency of applied linguists to advocate major changes in teaching strategies based on a theory of SLA without the support that would be derived from the long-term trialling of what is advocated. Moreover, it has shown that the applied linguistic literature of the day has shown a marked reluctance to submit to critical scrutiny the advocacy of new teaching strategies and has accepted largely without question the premature rejection of previously-used teaching options. Thus, the advocacy of a task-based instruction using a FoF whilst proscribing a FoFS, has been allowed to develop during the last decade largely unscathed by the critical scrutiny to which it is so vulnerable.

The findings of the eight-month study described in this chapter demonstrate that the students in this study, who had spent the previous two years in classes taught by a teacher who adopted a SCLT approach with added FoF during the period of the study, had produced students with good comprehension skills but who were unable to produce complete accurate utterances manifesting the internalisation of any grammatical system, thus casting serious doubt on the contention that learners can acquire parts of the grammar thanks solely to exposure to interactive CI.

Further, the results of the study demonstrated that the students benefiting from a FoFS targeting interrogative forms and adverb placement achieved dramatic improvement in oral performance in those areas whilst the students continuing to benefit from SCLT with a FoF, made no progress in oral performance in the two target areas.

There are two conclusions to be drawn from this and the other available evidence on the value of a FoFS. First, all future comparative research on the efficacy of form-focussed instruction must necessarily include an exponent of a FoFS approach. Second, given the evidence in favour of a FoFS and the less-than-compelling findings in favour of a FoF, it seems unjustified to advocate reforms which exclude a FoFS from the range of teaching strategies available to the teacher.

## Notes

1. Note that Long here refers to the 1999 conference presentation and not Norris & Ortega's 2000 published article.
2. The pretest, posttest 1 and posttest 2 for the adverbs corresponds to, respectively, posttest 2, posttest 3 and posttest 4 for the interrogatives.
3. One other study (Spada and Lightbown 1993) has endeavoured to compare the effects on the teaching of question forms to francophone Quebecers learning English at the same level as in the study reported on here. The study intended to compare the effects of a FoFS treatment to which was subjected the EG with a FoF used to teach the CG. Unfortunately, it transpired that the teacher of the CG "...frequently corrected students' use of question forms (as well as other grammatical errors) and in some cases reminded students of the metalinguistic information about question formation that she had provided in earlier classes" (p. 213). This disqualifies the CG treatment in this study as a FoF and, therefore, detracts from the reliability of the findings in terms of the comparison between the effectiveness of the two treatments.

**Appendix 1**

Name: ..... Date: .....

Class: ..... School: .....

Listen to the following passage and then answer the following questions on it. It will be read a second time to enable you to check your answers.

Yesterday morning, Peter got up at 9 and left the house at ten thirty. He was going to the local shops to do food shopping with the 20 dollars his mother had given him. On the way there, he met his three friends, John, Mary and Robert, who asked him to go and play soccer with them. He wanted to go with them but he could not because his mother was waiting for the food. But as he liked soccer so much, he quickly went to the shops, bought the food, took it home and then ran to join his friends at the local park. After the game, the three boys went to the Mall for a drink and french fries but Mary had to go home. Peter got home at one o'clock.

1. What time did Peter leave home?  
\_\_\_\_\_
2. Where was he going to?  
\_\_\_\_\_
3. Why was he going there?  
\_\_\_\_\_
4. Whom did Peter meet?  
\_\_\_\_\_
5. What did they want Peter to do?  
\_\_\_\_\_
6. What did Peter do first?  
\_\_\_\_\_
7. How many friends played soccer?  
\_\_\_\_\_
8. Where did they play?  
\_\_\_\_\_
9. Where did Peter go after playing soccer?  
\_\_\_\_\_
10. What did Peter's mother give him?  
\_\_\_\_\_
11. What did Mary eat after playing soccer?  
\_\_\_\_\_
12. How long was Peter absent from home?  
\_\_\_\_\_



**Appendix 2**

Name: ..... Date: .....

Class: ..... School: .....

Read the following sentences. If you think a sentence is written in correct English, write a capital C at the end. If you think there is a mistake in the sentence, rewrite it correctly.

1. Jean live in Quebec. ....
2. Do you like tennis? .....
3. Where are you born? .....
4. He do not likes chicken. ....
5. What are you go to do tonight? .....
6. Where your friend he live? .....
7. How long you live here? .....
8. We do not like school. ....
9. I go sometimes to the cinema. ....
10. He was born in Montreal. ....
11. My sister does not like meat. ....
12. Are you go to play chess today? .....
13. I often go to the cinema. ....
14. How many times your sister live here? .....
15. Where was your sister born? .....
16. Are you going to watch TV today? .....
17. I am go to play chess today. ....
18. Where does your friend live? .....
19. What did your friend do today? .....
20. Where did you went today? .....
21. Is Marie going to play soccer today? .....
22. How long has Jean lived here? .....
23. Is Pierre born in Quebec? .....
24. Where does your brother live? .....
25. Anne reads seldom the paper. ....
26. He goes never to the cinema. ....
27. How long have your brothers lived here? .....
28. What do your brother likes? .....
29. What is Jean going to do tomorrow? .....
30. Since when has she lived here? .....

## References

- Adams, R. and D. Chen  
1981 *The process of educational innovation: An international perspective.* London: Kogan Page in association with the UNESCO Press.
- Basturkmen, H., S. Loewen and R. Ellis  
2004 Teachers's stated Beliefs about Incidental Focus on Form and their Classroom Practices. *Applied Linguistics* 25: 243–270.
- Brookes, N.  
1964 *Language and Language Learning: Theory and Practice.* New York: Harcourt Brace.
- Brumfitt, C.  
1981 Notional syllabuses revisited. A response. *Applied Linguistics* 2: 90–92.
- Bruton, A.  
2002 From tasking purposes to purposing tasks. *ELT Journal* 56/3: 280–288.
- Carroll, S. and M. Swain  
1993 Explicit and implicit negative feedback: an empirical study of the learning of linguistic generalizations. *Studies in Second Language Acquisition* 15: 357–86.
- Catford, J. C.  
1998 Language Learning and Applied Linguistics: A historical sketch. *Language Learning* 48: 465–496.
- Celce-Murcia, M., Z. Dornei and S. Thurrell  
1997 Direct approaches in L2 instruction: A turning point in communicative language teaching. *TESOL Quarterly* 32: 141–152.
- Chastain, K. D.  
1970 A Methodological Study Comparing the Audio-lingual Habit Theory and the Cognitive Code-Learning Theory – a Continuation. *The Modern Language Journal* 54: 257–266.  
1988 *Developing Second-Language Skills: Theory and Practice.* New York: Harcourt Brace and Jovanovich
- Cook, V.  
1991 *Second Language Learning and Language Teaching.* New York: Edward Arnold.
- DeKeyser, R. M.  
1998 Beyond focus on form: Cognitive perspectives on learning and practising second language grammar. In C. Doughty and J. Williams (eds.), *Focus on Form in Classroom Language Acquisition*, 42–63. Cambridge: Cambridge University Press.
- Diller, K. C.  
1978 *The Language Teaching Controversy* Rowley MA: Newbury House.

- Doughty, C.  
2001 Cognitive underpinnings of focus on form. In P. Robinson (ed.), *Cognition and Second Language Instruction*. Cambridge: Cambridge University Press.
- Doughty, C. and J. Williams (Eds.)  
1998a *Focus on Form in Classroom Second Language Acquisition*. Cambridge: Cambridge University Press.  
1998b Pedagogical choices in focus on form. In C. Doughty and J. Williams (eds.), *Focus on Form in Classroom Language Acquisition*, 197–261. Cambridge: Cambridge University Press.
- Doughty, C. and E. Varela  
1998 Communicative Focus on Form. In C. Doughty and J. Williams (eds.), *Focus on Form in Classroom Language Acquisition*, 114–138. Cambridge: Cambridge University Press.
- Dulay, H., M. Burt and S. Krashen  
1982 *Language Two*. New York: Oxford University Press.
- Ellis, R.  
1994 *The study of Second Language Acquisition*. Oxford: Oxford University Press.
- Ellis, R., H. Basturkmen and S. Loewen  
2001 Pre-emptive Focus on Form in the ESL Classroom. *TESOL Quarterly* 35: 442–467.
- Erlam, R.  
2003 Evaluating the relative effectiveness of structured input and output-based instruction in foreign language learning: Results from an experimental study. *Studies in Second Language Acquisition* 25: 559–582.
- Fullan, M.  
1982 *The meaning of educational change*. New York: Teachers College Press.
- Gass, S. and L. Selinker (eds.)  
1983 *Language Transfer in Language Learning*. Rowley, Mass.: Newbury House.
- Germain, C.  
1993 *Evolution de l'enseignement des langues: 5000 ans d'histoire*. Paris: Hurtubise HMH, Ltée.
- Gouvernement du Québec  
1986 Guide pédagogique secondaire: anglais langue seconde. Direction générale des programmes.
- Harley, B.  
1989 Functional grammar in French immersion: A classroom experiment. *Applied Linguistics* 10: 321–51.
- Howatt, A. P. R.  
1984 *A History of English Language Teaching*. Oxford: Oxford University Press.

- Johnson, K.  
 2001 *An Introduction to Foreign Language Learning and Teaching*. Harlow: Longman.
- Kellerman, E. and M. Sharwood Smith  
 1986 *Crosslinguistic Influence in Second Language Acquisition*. Oxford: Pergamon Press.
- Krashen, S.  
 1985 *The Input Hypothesis: Issues and Implications* London: Longman
- Kupferberg, I. and E. Olshtain  
 1996 Explicit contrastive instruction facilitates the acquisition of L2 forms. *Language Awareness* 5: 149–165.
- Lantolf, J. P.  
 2000 A century of language teaching and research: Looking back and looking ahead. *The Modern Language Journal* 84: 467–471.
- Larsen-Freeman, D. and M. H. Long,  
 1991 *An introduction to second language acquisition theory and research*. London: Longman.
- Lightbown, P. M.  
 1991 What have we here? Some observations on the influence of instruction on L2 learning. In R. Phillipson, E. Kellerman, L. Selinker, M. Sharwood Smith and M. Swain (eds.), *Foreign/second language pedagogy research*, 197–212. Clevedon, Avon: Multilingual Matters.
- 1998 The importance of timing in focus on form. In C. Doughty and J. Williams (eds.), *Focus on Form in Classroom Second Language Acquisition*, 177–196. Cambridge: Cambridge University Press.
- 2000 Anniversary article: Classroom SLA research and second language teaching. *Applied Linguistics* 21: 431–462.
- 2002 The role of SLA research in L2 teaching: Reply to Sheen. *Applied Linguistics* 23 (4): 530–536.
- Lightbown, P. M. and N. Spada  
 1990 Focus on form and corrective feedback in communicative language teaching: Effects on second language learning. *Studies in Second Language Acquisition* 12: 429–448.
- Long, M. H.  
 1988 Instructed interlanguage development. In L. Beebe (ed.), *Issues in second language acquisition: Multiple perspectives*, 115–141. Rowley, Mass.: Newbury House.
- 1991 Focus on form: A design feature in language teaching methodology. In K. de Bot, R. Ginsberg and C. Kramsch (eds.), *Foreign language research in cross-cultural perspective*, 39–52. Amsterdam: John Benjamins
- 1994 Response to ‘A critical analysis of the advocacy of the task-based syllabus’ [*TESOL Quarterly* 28: 127–151]. *TESOL Quarterly* 28 (4): 782–790.

- 2000 Focus on form in task-based language teaching. In *Language policy and pedagogy: Essays in honor of A. Ronald Walton*. R. D. Lambert and E. Shohamy (eds.), 179–192. Philadelphia: John Benjamins.
- Long, M. H. and G. Crookes  
1992 Three approaches to task-based syllabus design. *TESOL Quarterly* 26: 27–56.
- Long, M. H. and P. Robinson  
1998 Focus on form: Theory, research and practice. In C. Doughty and J. Williams (eds.), *Focus on Form in Classroom Language Acquisition*, 15–41. Cambridge: Cambridge University Press.
- Lyster, R.  
1994 La négociation de la forme: Stratégie analytique en classe d'immersion. *Canadian Modern Language Review* 50: 446–65.  
1998 Negotiation of form, recasts, and explicit correction in relation to error types and learner repair in immersion classroom. *Language Learning* 48: 183–218.
- Markee, N.  
1993 The diffusion of innovation in language teaching. *Annual Review of Applied Linguistics* 13: 229–243.
- Meara, P.  
1998 Review of J. Coady and T. Huckin (eds.), 1997, *Second language vocabulary acquisition: A rationale for pedagogy*. *Applied Linguistics* 19: 289–292.
- Musumeci, D.  
1997 *Breaking tradition: An exploration of the historical relationship between theory and practice in second language teaching*. New York: McGraw Hill.
- Nassaji, H.  
1999 Towards integrating form-focused instruction and communicative interaction in the second language classroom: Some pedagogical possibilities. *The Canadian Modern Language Review* 55: 385–402.
- Norris, J. M. and L. Ortega.  
2000 Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning* 50: 417–528.
- Odlin, T.  
1989 *Language Transfer – Cross-linguistic Influence in Language Learning*. Cambridge: Cambridge University Press.
- Palmer, A.  
1992 Issues in evaluating input-based language teaching programs. In J. C. Alderson and A. Beretta (eds.), *Evaluating Second Language Education*, 141–166. Cambridge: Cambridge University Press.
- Phillipson, R., E. Kellerman, L. Selinker, M. Sharwood Smith and M. Swain (eds.)  
1991 *Foreign/Second Language Pedagogy Research*. Clevedon, PA: Multilingual Matters.

- Richards, J. C. and T. S. Rodgers  
 1986 *Approaches and methods in language teaching*. Cambridge: Cambridge University Press.
- Robinson, P.  
 1997 State of the Art: Research and Second Language Teaching. *The Language Teacher* 21: 7–16.
- Scott, V. M.  
 1989 An empirical study of explicit and implicit teaching strategies in French. *The Modern Language Journal* 73: 14–21.
- Sheen, R.  
 1993 Double standards in research selection and evaluation: The case of comparison of methods, instruction and task-work. *RELC Journal* 24: 1–28.  
 1994a A critical analysis of the advocacy of the task-based syllabus. *TESOL Quarterly* 28: 127–151.  
 1994a A rebuttal of the response of M. K. Long to ‘A critical analysis of the advocacy of the task-based syllabus’ [*TESOL Quarterly* 28: 127–151]. *TESOL Quarterly* 28 (4): 782–790–795.  
 1996 The advantage of exploiting contrastive analysis in teaching and learning a foreign language. *International Review of Applied Linguistics* 24: 183–197.  
 1999a A success story. *SPEAQ OUT* 27/4: 14–16.  
 2000 A critical review of Catherine Doughty and Jessica Williams (eds.), 1998, *Focus on form in classroom Second Language Acquisition*, (The Cambridge Applied Linguistics Series). Cambridge University Press: Cambridge, UK. 301 pages. *The Linguist List*. Internet publication. Pages: 25.  
 2003 Focus in form – a myth-in-the-making. *English Language Teaching Journal* 57: 225–233.  
 2004 A plenary too far... – a critique of VanPatten, B., 2004, The Fundamental Similarity Hypothesis. Opening plenary given at the IATEFL Conference in Liverpool, UK. *IATEFL Issues, October 2004*.  
 2005 Developmental sequences under the microscope. *Proceedings of The IATEFL Annual Conference 2004 in Liverpool, UK*.
- Skehan, P.  
 1996 A framework for the implementation of task-based instruction. *Applied Linguistics* 17: 38–62.  
 1998 *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.
- Smith, P. D. Jr.  
 1970 A comparison of the cognitive and audio-lingual approaches to foreign language instruction. *The Pennsylvania Foreign Language Project*. Philadelphia: The Center for Curriculum Development.

- Spada, N.  
1987 Relationships between instructional differences and learning outcomes: A process-product study of communicative language teaching. *Applied Linguistics* 18: 137–161.  
1997 Form-Focussed Instruction and Second Language Acquisition: A Review of Classroom and Laboratory Research. *Language Teaching* 30: 73–87.
- Spada, N. and P. M. Lighbown  
1993 Instruction and the development of questions in the L2 classroom. *Studies in Second Language Acquisition* 15: 205–221.  
2002 Second Language Acquisition. In N. Schmitt (ed.), *An Introduction to Applied Linguistics*, 115–132. New York: Oxford University Press.
- Spolsky, B.  
1989 *Conditions for second language learning*. Oxford: Oxford University Press.
- Stern, H. H.  
1992 *Issues and options in language teaching* Oxford: Oxford University Press.
- Strevens, P. D.  
1987 The nature of language teaching. In M. H. Long and J. C. Richards (eds.), *Methodology in TESOL: A book of readings*, 10–25. Boston, Mass.: Heinle and Heinle Publishers.
- Swan, M.  
1988 Manipulating and complementing content teaching to maximize second language learning. *TESL Canada Journal* 6: 68–83.  
forthc. Legislating by hypothesis: Focus on form and task-based learning. *Applied Linguistics*.
- Ur, P.  
1993 Teaching Grammar. In *A Communicative Methodology*, 139–145. Israeli Ministry of Education.
- Valette, R. M.  
1991 Proficiency and the prevention of fossilization – an editorial. *The Modern Language Journal* 75: 325–336.
- VanPatten, B. and C. Sanz  
1995 From input to output: Processing instruction and communicative task. In F. Eckman, D. Highland, P. Lee, J. Mileham and R. Weber (eds.), *SLA theory and pedagogy*, 169–185. Hillsdale, NJ: Lawrence Erlbaum.
- Von Elek, T. and M. Oskarsson  
1973 *Teaching Foreign Language Grammar to Adults: A comparative study*. Almqvist and Wiksell: Stockholm.

White, J.

- 1998 Getting the learners' attention. In C. Doughty and J. Williams (eds.), *Focus on Form in Classroom Second Language Acquisition*, 85–113. Cambridge: Cambridge University Press.
- 2001 How can teaching affect progress on a grammatical feature? *SPEAQ Annual Convention, 1–3 November, 2001*. St Hyacinthe, Quebec, Canada.



# **Instructed Second Language Vocabulary Learning: The fault in the ‘default hypothesis’**

*Batia Laufer*

The ‘default hypothesis’ of vocabulary acquisition claims that we acquire most words in our native language through exposure to language input, particularly written input, rather than by instruction. The paper examines the basic assumptions underlying the default hypothesis: the noticing assumption, the guessing ability assumption, the guessing-retention link assumption, the repeated exposures – retention link assumption, and the extrapolation assumption. It is argued and demonstrated that, in the context of instructed language learning, none of them can be taken for granted. Often second language learners do not notice unfamiliar words as unfamiliar, when they do, they often fail to infer their meaning from context, successful guessing does not necessarily lead to successful retention of meaning, repeated exposures require a flood of reading which is unrealistic in instructed language setting, results of vocabulary gains from short text cannot be extrapolated to larger quantities of reading. An alternative hypothesis is proposed, according to which, the main source of L2 vocabulary knowledge is likely to be Planned Lexical Instruction (PLI), which is in line with Focus on Form and Focus on Forms approaches. PLI ensures noticing, provides correct lexical information, and creates opportunities for forming and expanding knowledge through a variety of word focused activities.

## **Introduction**

How do people learn words in a second language? By deliberately committing them to memory? By using them in original contexts? By manipulating them in exercises devised by teachers? By encountering them in the language input?

The ‘default hypothesis’ of vocabulary acquisition, which was originally formulated for the native language, claims that we acquire most words through repeated exposures to language input, particularly to written input. It is justified in negative terms: the number of words that people know is

too vast to be accounted for by direct teaching of vocabulary (Nagy, Herman and Anderson 1985; Sternberg 1987; Nagy 1997). According to modest estimates, native speakers of English who are high school graduates know about 20000 word families, a figure which corresponds to about 32000 lexical items (Nation 1990) (1). Less modest estimates claim that the number is much larger, about 50000. Such figures seem to lend credibility to the 'vocabulary through reading position'. Furthermore, the default hypothesis is also supported by research showing that the largest vocabulary growth occurs when children reach literacy, during the primary and particularly secondary school years, when they are estimated to read approximately a million words of text a year (Anderson, Wilson and Fielding 1988). (2)

Even though the default hypothesis was developed on the basis of first language vocabulary acquisition, it has also found supporters among quite a few second language researchers and practitioners. Krashen and his colleagues have been particularly active in promoting the importance of reading for vocabulary acquisition, in naturalistic and instructed contexts (e.g. Krashen 1989; Dupuy and Krashen 1993; Cho and Krashen 1994). The claim that vocabulary in a second language is determined by reading has important pedagogical implications, particularly for instructed L2 learning. If reading is indeed the main source of vocabulary acquisition, then instead of word focussed practice, learners should be required to read as much as possible in and outside the classroom. But is reading the main source of L2 vocabulary? I will try to examine the default hypothesis in the light of conditions which are characteristic of instructed language learning. This type of learning refers to learning the second language as a subject, rather than learning school subjects via the medium of the second language. Instructed language learning takes place mainly in a classroom, since it is not the language of the country in which it is taught. Learners may be exposed to this language outside the classroom via television and computers, but this exposure is limited and cannot be expected to occur with all the learners to the same degree. Hence, the main source of second language learning is the classroom. The number of teaching hours amounts to several hours per week, depending on the type of school and the year of instruction (3).

### **Assumptions underlying the default hypothesis**

The claim that people acquire vocabulary through reading presupposes five assumptions on which this claim rests. I will refer to these as: the noticing assumption, the guessing ability assumption, the 'guessing-retention link'

assumption, the 'repeated exposures – retention link' assumption, and the extrapolation assumption.

*The noticing assumption*

On encountering an unfamiliar word, the reader notices it as a word s/he does not know. Gass (1988) refers to such an encounter as an apperceived input. Noticing or attention to the new material is an essential condition for learning (Schmidt 1994). Those who argue that some learning can be implicit, i.e. occur without attention to what is being learnt, state that this is so in the case of grammar, not vocabulary (Paradis 2000). Noticing a word as unfamiliar is a prerequisite for the next assumption.

*The guessing ability assumption*

On encountering and noticing an unfamiliar word, the learner decides to infer its meaning from context. Guessing strategy for reading comprehension has been advocated by some reading experts (Haastrup 1991; Schouten-van Parreren 1989; Elley 1989) and taught in first and second language courses. The basic assumption behind this recommendation is that people can use contextual clues successfully and will do so, particularly, if the word is deemed relevant to the message of the text.

*The 'guessing-retention link' assumption*

Learning the word's meaning, let alone meaning and the other word properties, implies more than comprehending it in a particular text during a reading activity. The meaning of word has to be retained in long term memory. In surveying the rate of learning from context, Nagy (1997) concludes that first language readers have one-in-twenty chance of retaining successful guesses.

*The 'repeated exposures-retention link' assumption*

If a word is not remembered after the learner's first exposure to it, additional encounters are needed in order to increase the probability of retaining it. But what is the optimal number of encounters needed for this? There is no definite answer to this question. Horst, Cobb and Meara (1998) suggest 8 exposures may be enough, Saragi, Nation and Meister (1978) suggest 12.

*The extrapolation assumption*

Even if one-in-twenty-chance of retaining a word yields a very small number of words, the cumulative gains over time may be quite remarkable if we extrapolate from this figure to a large number of texts and to longer periods of time. Nagy (1997) points out that if an average first language reader is exposed to a million words of text a year and 2%, i.e. 20,000 of these are unknown, then learning one word in 20 would amount to an annual gain of 1000 words. Using another type of extrapolation, researchers could claim the following: if it is demonstrated, in an experiment, that 5 words are learnt from one text, then 50 words will be learnt from 10 similar texts, 500 from 100 texts, and so on.

**The assumptions of the default hypothesis and instructed language learning.**

*The noticing assumption:*

*Do L2 learners understand how much they do not understand?*

Second language learners will not necessarily recognise unfamiliar words to be unfamiliar. Laufer and Yano (2001) investigated how accurately learners from three different countries could assess their understanding of words in text context. Learners self assessed their comprehension of a sample of words in the text and were subsequently tested on these words. The subjective evaluations were compared with the objective scores. The comparison showed that all learners over-estimated their understanding of words, often by over 60%.

For example, a learner reported that he understood 16 words out of the 20 words that were tested. However, when he was asked to translate the 20 words, he could translate correctly 10 words only. Hence his over-evaluation was 60% vis à vis his real knowledge. The over-evaluation was found to be similar for men and women. It was different, however, in the three different cultures. The most 'Western' group of subjects, the Israeli group, exhibited the largest mismatch between self perceived and real lexical understanding. The Japanese learners were the most modest ones, but they too over-evaluated themselves.

One reason for this over-evaluation may be that the learners who understand the overall message of the text do not pay attention to the precise meaning of individual words. Another reason has to do with learners' unawareness of an unfamiliar word because it is confused with another word.

For example, a student who has come across 'adapt' may think s/he is reading 'adopt' rather than an unfamiliar word which should be guessed or looked up. Additional examples of frequent confusions are 'comprehensive' with 'comprehensible', 'counsel' and 'council', 'embrace' and 'embarrass', and many other similar lexical forms, or 'synforms' (Laufer 1988). Homonyms, false cognates, words with a deceptive structure (e.g. shortcomings, outline) are also prone to confusions (see Laufer 1989 for the discussion of Deceptive Transparency). Though the phenomenon of confusing similar lexical forms occurs in second language learning in general, instructed language learners are more susceptible to such confusions. Because of the relative paucity of input, words are encountered less frequently than in the natural language learning context (4). Consequently, many words are somewhat familiar but not yet firmly established in the mental lexicon and prone to confusions with similarly looking or sounding words. Hence, the noticing assumption cannot be taken for granted since in many cases, on encountering a new word, the learner does not notice it as such.

*The guessing assumption:*

*Can L2 readers infer the meaning of unknown words from context?*

When a word is noticed as unfamiliar, the reader may try and infer its meaning from context. Yet not all contexts provide clues for unknown words. In the sentence "I saw an X last night", there is no way to know what X is. Some contexts can even be misleading. In the sentence "People were drinking, singing, laughing, brawling", most learners interpreted "brawling" as "having a good time". Sometimes clues are ignored when the reader thinks s/he understands the message, or when the correct meaning is not compatible with the learner's schema. In the sentence "This shows how tenuous the relationship between fathers and sons is, for it can be destroyed by society", the word "tenuous" was interpreted "strong" even though the clause "for...society" was a good clue for the correct meaning. Most importantly, clues may appear in words which themselves are unknown to the learner and are therefore unusable. (For a discussion of the problems with guessing, see Laufer 1997). L1 readers, who do not understand a small number of words in an otherwise comprehensible context, can make a good use of available clues. Foreign learners, on the other hand, whose lexical coverage, i.e. the number of known words in the text, is below 98% will have a considerable difficulty if they try to infer the meaning of an unknown word from context (Hirsh and Nation 1992). If the coverage is below 95%, guessing becomes mission almost impossible (Laufer 1997). If

instructed learners are expected to successfully guess from context, they should be reading at their optimal lexical coverage, between 95% and 98%. This can hardly be expected in a normal classroom context. It is hard, if not impossible, for a classroom teacher, to select material for extensive reading which has an optimal lexical coverage. To do so, s/he would have to analyse the lexis of the texts and conduct extensive testing of learners' lexical knowledge in order to establish learner-text lexical compatibility (5). Since materials selected for reading are not based on such an extensive analysis and comparison, we cannot assume that guessing unfamiliar words will be successful.

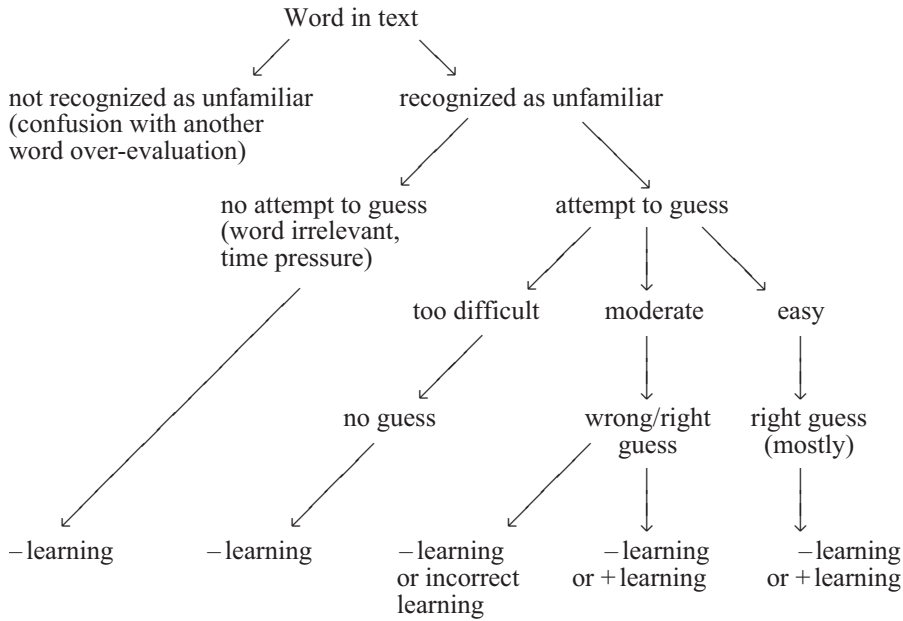
*The 'guessing-retention link' assumption:  
Do L2 learners retain the words they guess?*

When unfamiliar words in a text are deemed unimportant to the comprehension of the text, learners will often ignore them. Words deemed important will be attended to (Hulstijn et al 1996). Of these, some will be guessed incorrectly. If the incorrect meaning is retained, the word cannot be considered learned. Other words may be guessed correctly. Of these, some will be guessed without any difficulty, either because the context is particularly revealing, or because the context is moderately revealing but the learner has a very good guessing ability. Quick and easy guessing is indeed an asset in comprehension. However, this asset works against retention of the guessed words. In a controlled experiment, Mondria and Wit de Boer (1991) compared guessing and retention of words in what they called "pregnant" contexts, "moderately pregnant" and "non pregnant" contexts, i.e. contexts which provided sufficient clues for guessing, some clues and no clues respectively. They found that words which appeared in "pregnant" contexts, were guessed with ease, but were not remembered well.

A similar conclusion is expressed by Jacoby, Crail and Begg (1979) and Haastруп (1991), who argue that words which are guessed with some difficulty will be remembered better since difficulty leads to processing effort, which, in turn, creates a more distinctive memory trace. Guessing in moderately difficult contexts is a useful classroom activity and could be practiced during intensive reading. Yet when the learners are expected to read large quantities of text on their own and in a limited period of time, it is doubtful whether they will interrupt the flow of reading and invest time in "guessing with difficulty". Words which can be guessed with ease, without interfering with reading, will indeed be guessed. The problem is that most of them will not be retained.

Interim summary

The following chart sums up the 3 assumptions above. It demonstrates what happens when a learner comes across a new word in a text



What the chart shows is that there is very little chance for a word to be learned after the first exposure. If it is not recognized as unknown, there will be no attempt to guess it and consequently no learning. If it is noticed as unfamiliar, the learner may choose not to guess it, and consequently not to learn the word. If s/he want to guess it, guessing may be impossible or difficult due to reasons discussed earlier. In this case too, no learning of the word has occurred. In a moderately difficult context, the guess can be correct or incorrect. If it is incorrect, no learning has taken place. If it is correct, there is a chance that the word will be retained, and hence no learning has occurred. If guessing is very easy, the guess can be correct, but chances are that the word won't be remembered. As the chart shows, learning a word after encountering it in a text is contingent upon several conditions which may not prevail in a real reading activity. A valid counter-argument is that one exposure does not result in learning a word, but repeated exposures do. We will examine this argument in the next section.



*The 'repeated exposures-retention link' assumption:  
How much do learners need to read to 'pick up' a word?*

When learning from context occurs, it is usually a cumulative process. First encounters with a word may result in partial or vague understanding. Each additional exposure to the same word may enrich and strengthen the learner's knowledge of it. If, as pointed out earlier, about 10 exposures are necessary before the learner can recall or recognize the meaning of a word, the crucial question is how much reading has to be done to ensure these 10 repetitions of the new vocabulary.

Nation and Wang (1999) examined the lexis of graded readers. Graded readers at level 5 assume a knowledge of 1487 word families and introduce 459 new word families. Of these, only 108 occurred 10 times or more in the readers investigated at this level. The authors also calculated the number of running words which will ensure the average of 10 repetitions per word. The number was found to be 200,000. Thus, the meaning of 108 words will have a good chance to be learnt passively when a text of 200,000 words has been read. As the actual average text length at that level is 22,627 words, students have to read nine books to have a realistic chance to learn the meaning of 108 words. (In authentic non simplified texts, much larger quantities of text would have to be read to meet the same words). Furthermore, reading the above quantity of text has to be carried out when all the other conditions discussed before are fulfilled: learners understand 95%–98% of the text, notice the new words 10 times, attempt to guess them, the guessing conditions are favourable and the guess is, at least partially successful. If these ideal conditions are not fulfilled, fewer words will have a chance to be learnt after 10 exposures.

The most important factor which contributes to retention is meeting the words again before they have been forgotten. Therefore, according to Nation and Wang (1999), reading needs to be done at an intensive rate of around a book or two books per week. The question is whether such flood of reading can be implemented in a classroom setting with 2, 3, or even 5 hours of instruction per week. If these few hours have to suffice for teaching language structure and developing all 4 skills (listening, writing, speaking, reading), then flood of reading, inside or outside the classroom, is an unrealistic expectation. Thus, though repeated exposures to the same word are indeed related to its retention, the fourth assumption (repeated exposures-retention link) which underlies the default hypothesis cannot be taken for granted in instructed language context.



*The extrapolation assumption*

Experiments on vocabulary learning through reading report very small gains of 1–5 words from relatively short texts of 1000–7000 words; (Day et al 1991; Knight 1994; Hulstijn 1992; Pitts et al 1989; Paribakht and Wesche 1997). Slightly higher gains (6 words) are reported by Dupuy and Krashen (1993) but this study included the use of video in addition to reading and learners knew they would be tested. In Cho and Krashen (1994), the subject who engaged in pleasure reading without using a dictionary learnt 7 words from a booklet of 7,000 words.

It is tempting to assume that if, for example, 3 words are learnt after reading a text of 500 words, then 30 words will be learnt with subsequent reading of 5,000 words. But there is *no evidence that this extrapolation is correct*. A calculation like this does not take into account memory fatigue of the reader and forgetting that might occur during a prolonged reading activity. More realistic estimates of lexical learning through reading are those obtained from studies on word 'pick up' from long texts. In a study by Horst, Cobb and Meara (1998), learners read and listened to a novel, *Mayor of Casterbridge*, read by the teacher. The book, which has 21,000 running words, was read during 6 class sessions. It was found that the average vocabulary pick up was 5 words. Two years earlier, Lahav (1996) conducted a study of vocabulary learning from simplified readers. She tested students who read 4 readers, each one of about 20,000 words, and found an average learning rate of 3–4 words per book. A limitation of the two studies is in the number of words tested, 45 words in the first study and 50 per book in the second. A larger number of test items might have revealed some acquisition of additional words.

Horst and Meara (1999) tried to overcome this limitation by developing and testing a model which establishes growth rates in one round of learning (a matrix). These are then used to predict vocabulary gains in subsequent reading. They showed that the initial growth matrix predicted subsequent learning quite well. The authors, however, acknowledge the limitations of the study. The participant in the case study was a highly motivated and skilled language learner. He read one novel 8 times and knew he would be tested on the new words. Once he knew the events of the story, he concentrated on the unfamiliar words. Hence, his vocabulary learning was different from incidental vocabulary acquisition subsumed by the default hypothesis. Meara and Sanchez (2001) suggest that long term vocabulary development forecast could be made by using TPM (Transitional Probability Matrix). It is a tool which shows the probability of movement of words between several

states of knowledge and is derived from data collected on 2 occasions for the same learners (6). Yet this model rests on the assumption that learners can assess themselves accurately. Earlier in the paper, evidence was presented that this may not be the case. Furthermore, long term predictions about development of knowledge in general, and vocabulary in particular, must assume that the learning conditions responsible for this development will not change. One may well calculate the probability that some words will be acquired on the basis of the data collected on 2,3, or 5 occasions. But stating that this probability will not change in the future rests on the assumption that the learning conditions do not change either. This is not necessarily the case in any type of learning, let alone in instructed language learning. As learners move from one grade to another, different areas of syllabus and different language skills get emphasized. Adolescents, in particular, may change interests. For example, at one period of time, they may regularly engage in computer games which are rich in English as a foreign language. Later, such games are abandoned as different interests develop.

The kind of research conducted by Horst and Meara, and Meara and Sanchez may eventually solve the extrapolation problem. But future research will have to investigate large numbers of typical language learners, who are not exceptionally skilled and motivated, and who will read several different texts rather than one text several times. It will also have to incorporate the variable of imprecision in self reports and instability of learning conditions. Until then, vocabulary gains found in studies using single texts, particularly short texts, cannot be used as good predictors of incidental vocabulary acquisition.

### **An alternative hypothesis**

The default hypothesis is based on the assumption that people possess large vocabularies, too large to be accounted for by form-focused instruction. But instructed L2 learning is a different case. Foreign language learners have small vocabularies. Empirical data on learners' vocabulary size in different countries show that L2 learners know small amounts of words in spite of many hours of instruction (Laufer 1999). For example, in Japan EFL university learners know (passively) 2000–2300 word families after 800–1200 hours of instruction (Shillaw 1995; Barrow et al. 1999), Indonesian EFL university learners know 1200 word families after 900 hours of instruction (Nurweni and Read 1999). If we convert these figures into the average number of words learned per teaching hour (2–3 words per hour), it is not

inconceivable that learning occurred as a result of planned instruction rather than exposure to reading input. Hence I would like to propose an alternative hypothesis for vocabulary learning in instructed learning context:

*In view of the special conditions which obtain in instructed language learning context, the main source of L2 vocabulary knowledge is likely to be word focused classroom instruction.*

This hypothesis is in line with form focused instruction in general. Though most discussions of 'focus on form' have been related to grammar, Ellis (2001) rightly points out that 'form' involves more than grammar. I will therefore adopt Ellis' definition of form focused instruction as "attention to lexical forms and the meanings they realize, where words are treated as objects to be learned" (Ellis 2001: 13). Just as learners' attention can be directed to grammatical features in language input, so can it be drawn to specific lexical items. This can occur in a meaning based task, in response to a communicative need, or in a decontextualized vocabulary activity. An example of the former would be a reading task in which learners are required to use a dictionary to check the meaning of unfamiliar words. An example of the latter is completing a crossword puzzle, in which each target word is unrelated to the other words. In Long's (1991) terminology, the first example would be an instance of Focus on Form, while the second – of Focus on Forms, i.e. focus on linguistic features in their own right rather than as language elements in a communicative event.

Similarly to grammar, form focused lexical instruction can be incidental, or pre-planned (7). An example of the former would be a class discussion in which learners' attention is drawn to words which are unfamiliar, or partially familiar, to some learners. Even though an oral task can be designed by a teacher in advance and some of the key lexical items anticipated, additional words are bound to be needed by different participants. This task would be defined by Long as Focus on Form. A pre-planned lexical instruction, on the other hand, defies the original definition of Focus on Form, which advocates incidental focus only, but not the modified definition, for example that of Doughty and Williams (1998). An example is a reading task with words pre-selected by the teacher for instruction in class. The primary focus of the task is meaning, and words, though pre-selected, are necessary for the successful completion of the task. However, there is another type of a pre-planned lexical activity, which fits neither the original, nor the modified definition of Focus on Form, for example, blanks of isolated sentences to be completed with missing words which are provided in a list. This is not a

communicative activity, since the primary focus is on the missing words in disconnected sentences. Hence, it would be classified as a Focus on Forms activity.

The current discussions of grammar instruction suggest that structures are best learned when focused on in a meaning-oriented communicative context. In other words, Focus on Form, rather than Focus on Forms, is claimed to benefit the learner. The question we should ask here is whether the same applies to lexis and lexical learning. I am not familiar with empirical research designed to specifically compare vocabulary acquisition in the two conditions above, focus on form and focus on forms. (Quian's (1996) study showed that decontextualized learning fared better than contextualized meaning-based learning, implying that focus on forms was effective). It would not surprise me if future studies demonstrated the superiority of focus on forms in lexical acquisition, and supported earlier findings summarized by Nation as follows: "In fact, we know that words in isolation are retained very well indeed, both in large quantities and over long periods." (Nation 1982: 23) The reason for this is the nature of second language vocabulary, as opposed to the nature of grammatical structures. Unlike grammar, which is a closed system of rules, vocabulary is an open set, albeit of interrelated items. Except for the most frequent words that are learnt, new vocabulary will not necessarily appear in the input soon after initial exposure to it. This is true in the case of the first language, let alone a foreign language in the context of instructed learning, where input is limited. Similarly, if left to their own devices, L2 learners will often prefer to use simple, well known vocabulary, as long as the intended meaning can be conveyed (Laufer 1998). While learners can hardly avoid hearing, for example, the past simple in the input, or using it, they can manage quite well in L2 for some time before a communicative need arises for words like 'ubiquitous', 'vindication', 'fallacy'. Furthermore, grammatical structures are likely to appear in various communicative tasks, regardless of their content. Vocabulary, on the other hand, is content bound, and cannot be expected to naturally reappear in tasks that are thematically unrelated. Therefore, to ensure multiple exposures to specific words and their use, careful planning is required on the part of the teacher. This may take the form of selecting the vocabulary to focus on, and supplementing meaning oriented tasks with non-communicative, focus on forms type activities. It is my belief that in vocabulary teaching, there is value for focus on form, focus on forms, incidental and pre-planned instruction. Sheen (2000) suggested an umbrella term for all these types of instruction for grammar – PGI, Planned Grammar Instruction. I will use a similar term for vocabulary – PLI, Planned Lexical Instruction.

### **A plea for PLI**

How can PLI (planned lexical instruction) overcome the shortcomings of vocabulary-through-reading approach? Earlier, I argued that learners do not necessarily 'notice' unfamiliar words in the reading input. This is somewhat similar to not attending to grammatical forms when processing input for meaning. Yet according to Schmidt (1994), there is no learning without attention. PLI makes sure that when words are selected for learning, they are attended to, whether in a reading passage, or a specially designed activity. One of such activities can be inferring from context. However, the teacher will make sure that guesses are verified, incorrect guesses corrected, and word meaning supplied when clues are unavailable, unusable or misleading.

Earlier I pointed out that guessing new words does not necessarily lead to their retention. I also questioned the feasibility of flood of reading, and hence, of repeated exposures to words in instructed learning situation. One feature of PLI is providing additional exposures to new words in planned input, rather than leaving such exposure to mere chance. More importantly, PLI implies that words are practiced in word focused activities with various degrees of learner's 'involvement', as suggested by Laufer and Hulstijn (2001). According to their 'hypothesis of involvement', the efficacy of a task for learning does not depend on whether it is communicative or not, or whether it entails incidental or intentional learning, but on the involvement load of the task. When instruction is planned, teachers can assign tasks of various involvement loads, depending on word difficulty, word importance, and the degree of word knowledge aimed at.

In sum, PLI compensates for the relative paucity of input and a limited reoccurrence of words in instructed learning context. It also ensures noticing, provides correct lexical information, and creates opportunities for forming and expanding knowledge through a variety of word focused activities.

### **Conclusion**

In this paper, I examined the basic assumptions underlying the hypothesis that most vocabulary in L2 is acquired incidentally from reading. These are the noticing assumption, the guessing ability assumption, the guessing-retention link assumption, the repeated exposures – retention link assumption, and the extrapolation assumption. I tried to show that in the context of instructed language learning, none of them can be taken for granted. On

seeing a new word, the learner does not necessarily 'notice' it, i.e. does not recognise it as an unfamiliar word, either because of a confusion with another word, or because of the tendency to overestimate one's understanding of words in text context. Noticing a word as new does not assure success in inferring its meaning. The most important cause of failure is insufficient lexical coverage of the text. Successful guessing does not necessarily lead to successful retention of meaning. Repeated exposures to the same word are indeed related to its retention, but to ensure repetitions of the same vocabulary, a 'flood of reading' is required, which is very hard to implement in classroom instruction. Experiments report very small vocabulary gains from short and long texts, and these results cannot be extrapolated to larger quantities of reading.

In view of the logical fallacy of the default hypothesis for instructed learning context. I proposed an alternative hypothesis, according to which, the main source of L2 vocabulary knowledge was likely to be word focused classroom instruction. I also argued that, in vocabulary instruction, there was value both in the Focus on Form, and in Focus on Forms approaches. Consequently, I suggested an umbrella term PLI (planned lexical instruction) for the kind of instruction which incorporated the two approaches and included both incidental and pre-planned vocabulary teaching. The value of PLI for instructed learning context lies in ensuring noticing, providing correct lexical information, repeated exposures to words, and ample opportunities for creating and expanding knowledge through a variety of word focused activities.

I am not arguing against the importance of reading for vocabulary learning. Reading can sometimes be the source of initial knowledge of words, it can help to expand the knowledge of already familiar words, or reinforce the memory of words not yet firmly established in the lexicon. And I am certainly not arguing against the educational value of reading activities. What I have tried to show is that, in instructed foreign language context, the main source of vocabulary knowledge is likely to be not reading, but Planned Lexical Instruction.

## Notes

1. A word family consists of a word, its inflections, and its common derivatives. A lexical item is a unit of meaning and its inflections.
2. Since these figures were obtained by written tests the results apply to literate native speakers only.
3. In Israel, for example, the number of hours devoted to English is usually 4–5 hours a week.
4. The input is impoverished if compared to naturalistic acquisition of learners who are regularly exposed to L2 and use it in educational and work contexts. Adult immigrants who surround themselves by speakers of their L1 and receive therefore impoverished L2 input are a different case. Their chances of meeting the same words repeatedly may not be different from those of instructed learners.
5. Analysis of Lexis in a text can be performed using special computer programmes, like VocabProfile, or Range (available from Nation's website <http://www.vuw.ac.nz/lals>). A good vocabulary size tests that can be used to estimate vocabulary size is the newly validated Levels Test (Schmitt, Schmitt and Clapham 2001)
6. Meara and Sanchez suggest at least 3 states of word knowledge: words not known are deemed to be in state 1; words which the testee is unsure about are in state 2; words which are reported as known by the testee are deemed to be in state 3. The model can be expanded by differentiating state 2 into different types of uncertainty.
7. Ellis (2001) distinguishes between planned and incidental form focused instruction. I chose the term 'pre-planned' to refer to teacher's intention to focus on a specific vocabulary in a lesson, or a task. I borrowed the term 'planned', which will be used later, from Sheen (2000) to refer to the teacher's decision to make use of all types of focused vocabulary instruction (pre-planned, incidental, direct and indirect)

## References

- Anderson, R. C., P. Wilson and L. Fielding  
1988 Growth in reading and how children spend their time outside of school. *Reading Research Quarterly* 23: 285–303.
- Barrow J, Y. Nakashimi and H. Ishino  
1999 Assessing Japanese College students' vocabulary knowledge with a self-checking familiarity survey. *System* 27: 223–247.
- Cho, K.-S. and S. Krashen  
1994 Acquisition of vocabulary from the Sweet Valley Kids Series: adult ESL acquisition. *Journal of Reading* 37: 662–667.



- Day, R.R., C. Omura and M. Hiramatsu  
 1991 Incidental EFL vocabulary learning and reading. *Reading in a Foreign Language* 8: 689–696.
- Doughty, C. and J. Williams  
 1998 Pedagogical choices in focus on form. In *Focus on Form in Classroom Second Language Acquisition*, Doughty, C. and J. Williams (eds.), 197–262. Cambridge: CUP.
- Dupuy, B. and S. Krashen  
 1993 Incidental vocabulary acquisition in French as a foreign language. *Applied Language Learning* 4: 55–63.
- Elley, W.B.  
 1989 Vocabulary acquisition from listening to stories. *Reading Research Quarterly* 24: 174–187.
- Ellis, R.  
 2001 Investigating form-focused instruction. *Language Learning* 51, supplement 1: 1–46.
- Gass, S.M.  
 1988 Second language vocabulary acquisition. *Annual Review of Applied Linguistics* 9: 92–106.
- Haastrup, K.  
 1991 *Lexical inferencing procedures or talking about words*. Tübingen: Gunter Narr.
- Hirsh, D. and P. Nation.  
 1992 What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language* 8: 689–696.
- Horst, M., T. Cobb and P. Meara  
 1998 Beyond a clockwork orange: acquiring second language vocabulary through reading. *Reading in a Foreign Language* 11: 207–223.
- Horst, M. and P. Meara  
 1999 Test of a model for predicting second language lexical growth through reading. *The Canadian Modern Language Review* 56: 308–328.
- Hulstijn, J. H.  
 1992 Retention of inferred and given word meanings: experiments in incidental vocabulary learning. In *Vocabulary and Applied Linguistics*, P. J. L. Arnaud and H. Bejoint (eds.) London: MacMillan.
- Hulstijn, J. H., M. Hollander and T. Greidanus  
 1996 Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *Modern Language Journal* 80: 327–339.
- Jacoby, L. L., F. J. M. Craik and J. Begg  
 1979 Effects of decision difficulty on recognition and recall. *Journal of Verbal Learning and Verbal Behaviour* 18: 585–600.



- Knight, S. M.  
1994 Dictionary use while reading: the effects on comprehension and vocabulary acquisition for students of different verbal abilities. *Modern Language Journal* 78: 285–299.
- Krashen, S.  
1989 We acquire vocabulary and spelling by reading: additional evidence for the input hypothesis. *Modern Language Journal* 73: 445–464.
- Lahav, J.  
1996 Vocabulary Acquisition Through Extensive Reading: 'Incidental' and Controlled Acquisition Compared. MA diss., University of Haifa.
- Laufer, B.  
1988 The concept of 'synforms' (similar lexical forms) in L2 vocabulary acquisition. *Language and Education* 2: 113–132.  
1989 A factor of difficulty in vocabulary learning: deceptive transparency. *AILA Review 6 (Vocabulary Acquisition)*: 10–20.  
1997 The lexical plight in second language reading: words you don't know, words you think you know and words you can't guess. In *Second Language Vocabulary Acquisition: a Rationale for Pedagogy*, J. Coady and T. Huckin (eds.), 20–34. Cambridge: Cambridge University Press.  
1998 The development of passive and active vocabulary in a second language: same or different? *Applied Linguistics* 19: 255–271.  
1999 Task effect on instructed vocabulary learning: the hypothesis of 'involvement'. Key-note address presented at AILA Congress. Tokyo.
- Laufer, B. and J. Hulstijn  
2001 Incidental vocabulary acquisition in a second language: the construct of task-induced involvement. *Applied Linguistics* 22: 1–26.
- Long, M.  
1991 Focus on form: a design feature in language teaching methodology. In *Foreign Language Research in Cross Cultural Perspective*, K. de Bot, R. Ginsberg and C. Kramsh (eds.), 39–52. Amsterdam: John Benjamins.
- Meara, P. and I. R. Sanchez  
2001 A methodology for evaluating the effectiveness of vocabulary treatments. In *Reflections on Language and Language Learning. In honour of Arthur van Essen*, M. Bax and J.-W. Zwart (eds.), 267–278.
- Mondria, J.R. and M Wit-de Boer  
1991 The effects of contextual richness on the guessability and the retention of words in a foreign language. *Applied Linguistics*. 12: 249–267.
- Nagy, W.E.  
1997 On the role of context in first- and second-language vocabulary learning. In *Vocabulary Description, Acquisition and Pedagogy*, N. Schmitt and M. McCarthy (eds.), 64–83. Cambridge: Cambridge University Press.

- Nagy, W. E., P. A. Herman and P. C. Anderson  
1985 Learning words from context. *Reading Research Quarterly* 20: 233–253.
- Nation, I. S. P.  
1982 Beginning to learn foreign language vocabulary: a review of the research. *RELC Journal* 13: 14–36.  
1990 *Teaching and Learning Vocabulary*. New York: Newbury House.
- Nation, P. and K. Wang Ming-tzu  
1999 Graded readers and vocabulary. *Reading in a Foreign Language* 12: 355–379.
- Nurweni, A. and J. Read  
1999 The English vocabulary knowledge of Indonesian University students. *English for Specific Purposes* 18: 161–175.
- Paradis, M.  
2000 Compétence linguistique et connaissance metalinguistique en fonction de différentes formes d'enseignement: apport de neurosciences. Paper presented at the Instructed Second Language Learning Colloquium. Vrije Universiteit Brussels.
- Paribakht, T. S. and M. Wesche  
1997 Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. In *Second Language Vocabulary Acquisition. A rationale for Pedagogy*, J. Coady and T. Huckin (eds.), 174–200 Cambridge: Cambridge University Press.
- Pitts, M., H. White and S. Krashen  
1989 Acquiring second language vocabulary through reading: a replication of the Clockwork Orange study using second language acquirers. *Reading in a Foreign Language* 5: 271–275.
- Quian, D. D.  
1996 ESL vocabulary acquisition: contextualization and decontextualization. *The Canadian Modern Language Review* 53: 120–142.
- Saragi, T., I. S. P. Nation and G. F. Meister  
1978 Vocabulary learning and reading. *System* 6: 72–28.
- Schouten-van Parreren, C.  
1989 Vocabulary learning through reading: which conditions should be met when presenting words in texts? *AILA Review 6 (Vocabulary Acquisition)*: 75–85.
- Schmidt, R.  
1994 Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review* 11: 11–26.
- Schmitt, N., D. Schmitt and C. Clapham  
2001 Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing* 18: 55–88.

- Sheen, R.  
2000 A response to Nassaji's 'Towards integrating form-focused instruction and communicative interaction in the second language classroom' in light of issues related to focus on form. *The Canadian Modern Language Review* 56: 496–513.
- Shillaw, J.  
1995 Using a word list as a focus for vocabulary learning. *The Language Teacher (JALT)* 19/2: 58–59.
- Sternberg, R. J.  
1987 Most vocabulary is learnt from context. In *The Nature of Vocabulary Acquisition* 89–105, M.G. McKeown and M.E. Curtis (eds.), Hillsdale, NJ: Erlbaum.



**Investigating the role and effects of  
interaction and Communication-Focused  
Instruction**



# **Negative feedback and learner uptake in analytic foreign language teaching**

*Katja Lochtmann*

Foreign language teaching (FLT) can be discussed in terms of the relationship between the concepts of “experience” and “analysis” (Stern 1990, 1992). In analytic FLT contexts the main focus is on the form(s) of the foreign language. Typical for analytic FLT are discrete point presentation along with extensive feedback on formal error. Experiential FLT mainly focuses on the meaning conveyed in the foreign language. However, several studies of experiential settings have shown that meaningful interaction alone does not allow students to achieve high levels of accuracy (Swain 1985; Lyster 1994). It is argued that foreign language students should “notice the gap” (Schmidt and Frota 1986) between their interlanguage and the target language and should be pushed to produce comprehensible and accurate output as well (Swain 1998). Negative feedback in the form of recasts (Doughty 2001) or negotiations of form (Lyster and Ranta 1997) could provide such “noticing” and/or “comprehensible output producing” opportunities.

On the basis of their survey of the literature on recasts in language learning Nicolas, Lightbown and Spada (2001: 720) argue that this type of negative feedback appears “to be most effective in contexts where it is clear to the learner that the recast is a reaction to the accuracy of the form, not the content, of the original utterance”. The question is how different types of negative feedback can be accounted for in terms of noticing the feedback in an analytic FLT setting (German as a foreign language in Belgium), where learners are generally supposed to pay attention to form(s) in the first place.

## **1. Introduction**

All foreign language teachers provide their learners with negative feedback on formal error (in this paper also referred to as “corrective feedback”). The way in which they do this depends on a number of parameters such as the teacher’s personality or attitude towards corrective feedback (Cathcart and

Olsen 1976; Nystrom 1983; Kleppin and Königs 1991; Lochtman 2002b), the learners' age and/or prior knowledge of the foreign language (see Kleppin and Königs 1991). This paper focuses on the foreign language teaching (FLT) context. According to Stern (1990, 1992) FLT can be discussed in terms of the relationship between the concepts "experience" and "analysis". Experiential and analytic FLT techniques can be understood as two ends of a continuum. While experiential FLT mainly focuses on the content or the message conveyed in the foreign language, analytic FLT focuses on the form of the foreign language or the medium itself. We assume that analytic FLT contexts differ from experiential FLT contexts with regard to feedback on error.

The issue of whether errors should be corrected or not has been investigated under the rubric of "focus on form" in classroom second language acquisition. According to Long (1991: 45–46) "Focus on form [...] overtly draws students' attention to linguistic elements as they arise incidentally in lessons whose overriding focus is on meaning or communication". "Focus on form" is often differentiated from the concept of "focus on forms", which "always entails isolation or extraction of linguistic features from context or from communicative activity" (Doughty and Williams 1998: 3). Whether oral corrective feedback is part of focus on form versus focus on forms FLT activities seems to depend on the instructional context of foreign language acquisition (FLA), i.e. experiential vs. analytic FLT.

One feature that has often been considered essential for FLT pedagogy is negotiated classroom interaction or negotiation of meaning (Ellis 1994, 1997; Long 1983, 1996). Negative feedback on formal error is regarded as an analytic teaching strategy (Lightbown and Spada 1990; Lyster and Ranta 1997). Interestingly, corrective feedback, as it often serves as the starting-point for negotiated interaction in the foreign language classroom (see Ellis 1994; Long 1983, 1996; Lyster and Ranta 1997) can thus also be regarded as part of experiential and more meaning-focused teaching activities. Negotiated interaction between the teacher and the learner in the foreign language classroom (FLC) is "characterised by interactional modifications such as comprehension checks and requests for clarification" (Ellis 1994: 716), which serve the purpose of solving problems in understanding and communication breakdowns and repairing language errors. Negotiation of meaning in getting your message across is seen as a feature of real language use, because it is both a characteristic of mother and child dyads in first language acquisition and of native-nonnative speakers' dyads in true conversation (Long 1983, 1996). This is one of the reasons why negotiation in the FLC is considered an opportunity for foreign language acquisition. In order to foster such



opportunities, some studies suggest that in FLCs pupils should be provided with comprehensible input (Krashen 1985) and should actively take part in meaningful interaction (Long 1983). Several studies of the experiential setting of immersion programmes, however, have shown that comprehensible input in meaningful interaction on its own is not enough for students to achieve high levels of accuracy in their second language (e.g. Swain 1985; Lyster 1994), and they propose that some focus on form is also required in order for students to improve their accuracy. Foreign language students should “notice the gap” between their IL and the TL (Schmidt and Frota 1986; Swain 1998; Doughty 2001) and should to a certain extent be pushed to produce comprehensible and accurate output as well (e.g. Swain 1985, 1995, 1998; Lyster 1994; Lyster and Ranta 1997). As observed by Schmidt (1995: 20) in the “noticing hypothesis”: “what learners notice in input becomes intake for learning”. Oral corrective feedback may be able to provide such noticing and/or comprehensible output producing opportunities.

According to Lyster and Ranta (1997), for example, negative feedback in the form of negotiation of form would meet these conditions. Starting from van Lier’s (1988) terminological distinction between conversational and didactic repair, Lyster and Ranta (1997) make a distinction between a conversational and a didactic function of negotiated interaction. The conversational function refers to negotiation of meaning, the didactic function involves what they call negotiation of form, which in turn refers to corrective feedback “that encourages self-repair involving accuracy and precision and not merely comprehensibility” (Lyster and Ranta 1997: 43). They argue that negotiations of form can only occur in interactions in which the teacher initiates a correction move and the students have the opportunity to correct themselves in the course of meaningful interaction. In keeping with Swain’s (1985) notion of comprehensible output, this should result in more accurate language use by the foreign language learners, because in this way learners are more likely to notice the feedback and will be pushed to produce comprehensible output as well. Lyster and Ranta (1997: 49) link the noticing issue to the presence or absence of uptake which “refers to a student’s utterance which immediately follows the teacher’s feedback and which constitutes a reaction in some way to the teacher’s intention to draw attention to some aspect of the student’s initial utterance”.

A few studies –apart from Lyster and Ranta’s (1997) study- have investigated the question of whether the learners have actually noticed the feedback in the input by looking at the learners’ reactions to the feedback. Chaudron (1986), for example, used a discourse analytic model in order to describe and analyse the correction behaviour of three (grades 8 and 9) immersion

teachers in Canada. He found that in this experiential setting the teachers seemed to prefer correction moves that initiated the pupils' self-correction. Only 39% of the correction moves in his study resulted in correct uptake. In his analysis Chaudron included both negative feedback on formal error and on content, however. Furthermore, most of the feedback he observed was on content rather than on formal error.

Kleppin and Königs (1991) and Havranek (1996, 1997, 1999) have investigated negative feedback on formal error in a more analytic FLT context. Kleppin and Königs' (1991) study was conducted in Spanish and Italian FLCs in German secondary schools. On the basis of classroom observations they concluded that correction behaviour was teacher-dependent. Interestingly, they found that the correction moves that initiated self-corrections were (what they call) unsuccessful, almost always resulting in needs-repairs. Havranek (1996, 1997, 1999) did an ethnographic study on oral corrective feedback in English FLCs in Austrian secondary schools and at university level. In order to investigate whether the learners had noticed the feedback she used analyses of classroom recordings, questionnaires and retrospection. She found that after a lesson only one third of the students could remember that they had been corrected. Especially recasts were not remembered. Explicit corrections or corrections with repetition of the correct answer appeared to be more salient. Interestingly, the students did not seem to interpret the initiations to self-correction as correction moves by the teacher.

The few studies discussed above show that the role of correction moves that initiate self-correction – which may lead to negotiation of form – is an issue for further investigation. Unfortunately, the results of these three studies cannot be compared because completely different methods of analysis were used (see also Norris and Ortega 2000).

## **2. Negative feedback in foreign language classroom interaction: recasts**

One condition for oral corrective feedback to foster FLA appears to be that it should be embedded in the process of “negotiation of/for meaning” and, in this way, does not interrupt the flow of discourse (Long 1996). However, interruptive explicit error correction hardly ever occurs in conversations outside an instructional FLA context (Long 1996). This led researchers to look for features of the interaction that include negative feedback in less direct ways (see e.g. Gordon 1990; Farrar 1992; Lightbown 2001). Long (1996: 429) has characterised such implicit negative feedback as “recasts”, which occur “in the form of implicit correction immediately following an

ungrammatical learner utterance". With regard to FLT a recast could be defined as "the teacher's reformulation of all or part of a student's utterance, minus the error" (Lyster and Ranta 1997: 48).

e.g. Learner: *Ich wünsche mir neuen Fussballschühe.*  
(‘I would like to have new football boots.’)  
teacher: *Aha, neue Fussballschuhe.*  
(‘Aha, new football boots.’) (from Lochtman 2000)

The psycholinguistic explanation for what happens in such recasts is that FL learners make an immediate cognitive comparison between their own erroneous utterance and the target language, recast by the discourse partner (Saxton 1997; Doughty and Varela 1998; Long et al. 1998; Mackey and Philp 1998; Doughty 2001). In order to be able to make such a cognitive comparison, it is commonly assumed that learners need to notice the feedback in the input. An issue which is still much debated, however, is whether recasts are a salient enough type of corrective feedback (Lyster 1998a; Lightbown 2001). By using recasts, teachers correct the learners' errors themselves, so that the learners are not explicitly pushed to produce repaired output.

According to Long (1996) the favourable circumstances (or psycholinguistic principles) for recasts to foster second language acquisition should be more clearly specified. Long (1996: 452) points out that recasts must also be semantically related to the learner's utterance. In FLT recasts should be embedded in semantically contingent speech which consists of "utterances by a competent speaker [the teacher], such as repetitions, extensions, reformulations, rephrasings and expansions". In this way Long once more stresses the importance of focussing on meaning at the moment formal errors are being corrected.

It appears that these conditions can only be met in experiential foreign language classrooms because in analytic classrooms, recasts hardly ever occur in semantically related or contingent speech. In analytic classrooms the recasts tend NOT to be "triggered by an error occurring as part of learners' attempt to communicate". Instead, they occur during activities like grammar exercises, when the attention of the interlocutors is "focused on an isolated target form for the explicit purpose of language learning [and] not in spontaneous conversation when the interlocutors are primarily focused on meaning" (Long 1996: 441).

While some researchers (Lyster 1998a, 1998b; Lightbown 2001) claim that in experiential FLT recasts may be too implicit to be useful in classroom

FLA because the learners do not seem to notice they are being corrected and their output shows no evidence that they have made a cognitive comparison, others (Long 1996; Doughty 2001) suggest that, when consistent with psycholinguistic principles mentioned above, recasts can deliver negative feedback favouring cognitive insight involving the noticing of inter-language vs. target language differences. However, on the basis of their survey of the literature on recasts in language learning Nicolas, Lightbown and Spada (2001: 720) argue that this type of negative feedback appears “to be most effective in contexts where it is clear to the learner that the recast is a reaction to the accuracy of the form, not the content, of the original utterance”.

The discussion above leads us to the following research questions:

- 1) What kinds of oral corrective feedback characterise analytic FLT where the focus is on forms?
- 2) How can these different kinds of oral corrective feedback be accounted for in terms of noticing the feedback?

### **3. German as a foreign language (GFL) in Flanders: analytic FLT**

German is the third foreign language in Flanders, after French and English, the L1 being Dutch. Foreign language teaching techniques in the Belgian secondary school system have always been mainly analytic, in the sense that “they have based themselves on some kind of analysis of the language” (Stern 1990: 94), often with an emphasis on grammar and error correction (see Lochtmán 1997). The language is presented in a relatively decontextualized form and the teacher and the learners are fully aware that they are practising the foreign language and are “simply displaying their command of the code” (Stern 1992: 302). This is also linked to the fact that the foreign language is both the medium and the subject of instruction. Furthermore, the concept of communicative language teaching appears to have been mainly defined as “teaching about communication” instead of “teaching through communication”. Teaching about communication entails “the analysis of communication through the study of speech acts, discourse analysis, and sociolinguistics” (Stern 1990: 94) and should therefore be situated on the analytic end of an experiential/analytic FLT continuum.

The GFL study reported on in this paper is classroom-based. Correction behaviour was not prescribed to the teachers, who were selected because of

their willingness to participate in this experiment. They were asked to act and correct as they normally did during their lessons. That they did not change their teaching method was verified on the basis of previous pilot study recordings of some of their lessons, when they did not know their correction behaviour was being observed.

The lessons were focused on forms and consisted of text comprehension and grammar exercises. The instructional materials targeted on plural formation in German and were treated as part of the regular teaching programme. The teaching activities were the same for the three teachers. First the teacher and the students read a short text of about 100 words aloud. Approximately one fifth of those words were plural forms. The teachers used the text as a starting point for the explanation of the plural formation in German (discrete point presentation). The students asked questions (mostly in the L1) and unknown or new vocabulary was translated into the L1. The students had to read the text aloud several times as a reading and pronunciation exercise. The grammar exercises that followed consisted of isolated sentences where plural forms had to be used. Some of the exercises (2 out of 6) were thematically linked to the text.

The teachers were unaware of the research focus related to corrective feedback, and thought the study focused on the acquisition of German plurals. The database analysed included 12 lessons totalling 600 minutes or 10 hours taught by three teachers. The recordings were made at the beginning of the fifth year of three secondary schools in Flanders. The pupils were all 15 or 16 years of age and should be considered relative beginners, since they only had had two hours a week of German in the previous school year.

For the purpose of the present study, the teachers' actual correction behaviour is described and analysed in the same way as in Lyster and Ranta (1997). Lyster and Ranta's (1997) data were derived from an observational study of French immersion classrooms in the Montreal area. This is a completely different setting, where the focus of the language lessons is on content or the message instead of on form or medium. This technique has therefore also been referred to as content-based language teaching, the content being another school subject (e.g. geography or history) and not the foreign language itself. Since it is hypothesized that some types of negative feedback might be more effective in contexts where it is clear to the learner that the recast is a reaction to the accuracy of the form, not the content, of the original utterance (Nicolas, Lightbown and Spada 2001), the findings of Lyster and Ranta's study are tentatively compared to those of the present study in order to see to what extent the learners might have noticed the feedback in a focus on forms environment.

#### 4. Feedback types and learner uptake

The following is a schematic representation of the different correction moves (see Henrici and Herlemann 1986; Lochtman 2000). The unit of analysis is the correction move initiated by the teacher. For a correction to take place there must first be an error or at least an utterance that according to the teacher should be corrected. The error is referred to as E. Following E there has to be a correction move C. Without these two elements there is no correction, therefore:

$$E \rightarrow C$$

The correction move can be a teacher initiated teacher correction (CT) or a teacher initiated self- or pupil-correction (CP), thus:

$$E \rightarrow CT$$

$$E \rightarrow CP$$

For a correction move to become negotiation of form there must be a student's response to the teacher's corrective feedback. This response is called uptake (U). Uptake is necessary for a correction move to be negotiation of form, but does not always occur. Therefore the final definition of a correction move is:

$$E \rightarrow CT \rightarrow (U)$$

$$E \rightarrow CP \rightarrow (U)$$

and of negotiation of form is:

$$E \rightarrow CP \rightarrow U$$

All negotiations of form are correction moves, but not the other way round. Note that only the CPs, i.e. the teacher initiated correction sequences that encourage self-correction, can be taken into account as negotiations of form. When there is learner uptake after a CT, i.e. when the teachers themselves corrected the student's error, the uptake is considered to be a repetition or echo of the teacher's utterance. In this case no negotiation has taken place.

The two different feedback types can be further subdivided (see also Lyster and Ranta 1997). The first feedback type, CT, comprises two moves:

- CT: 1. explicit corrections  
2. recasts





- Examples: E: "He goed home"
- \* clarification request: "Pardon me?"
  - \* metalinguistic feedback: "No, can you find the error?"
  - \* elicitation: "He \_?"
  - \* repetition: "He goed?"

Concerning learner uptake, Lyster and Ranta (1997) made a distinction between correct or successful uptake, referred to as "repair" and incorrect or otherwise unsuccessful uptake, referred to as "needs-repair". The category of "needs-repair" also includes student acknowledgements, such as "yes", "aha" or "oh", with which they just admit that they have made an error. Other examples of "needs-repairs" are student utterances with the same or a different error, hesitations etc. When there is no learner uptake, there is either topic continuation or the teacher might use a type of negative feedback once more to have the students repeat the correct answer or to have them correct themselves.

## 5. Results

In the 600 minutes of GFL classroom recordings, 394 correction sequences were identified for the three teachers together. 90% of the student turns which contained at least one error or were still in need of repair received negative feedback from the teacher. This contrasts, for example, with Lyster and Ranta's study, where only 62% of errors received feedback. Table 1 shows the distribution of the six feedback types in the GFL classrooms compared to the findings in Lyster and Ranta's study. To facilitate comparison the frequencies have been converted to percentages.

In the GFL classrooms one third (30.5%) of the teacher initiated teacher corrections (CTs) were recasts (Table 1). This is a major difference with regard to Lyster and Ranta's study, where this type of negative feedback comprises more than half of all the teacher correction moves. Furthermore, the GFL teachers made a high use of metalinguistic feedback (23.9%) and elicitations (30.2%), both correction types that provide the learners with the opportunity to correct themselves (CPs). These findings contrast with the few correction moves with metalinguistic feedback (8%) and elicitations (14%) used by the French immersion teachers in Lyster and Ranta's study. Surprisingly, clarification requests do not seem to belong to the GFL teachers' correction repertoire (1.8%). In experiential teaching the clarification requests occur more frequently (11%) than in the GFL study. Since



Table 1. The distribution of the different Feedback types as % of the total number of feedback moves

	Frequency	GFL (Flanders)	Lyster & Ranta (1997)
CT:			
1. Explicit corrections	54	13.7 %	7 %
2. Recasts	120	30.5 %	55 %
CP:			
3. Clarification requests	7	1.8 %	11 %
4. Metalinguistic feedback	94	23.9 %	8 %
5. Elicitations	119	30.2 %	14 %
6. Repetitions	0	0 %	5 %
		(n=394)	(n=687)

no repetitions of learner errors were found in the GFL study and only 5% of the correction moves in Lyster and Ranta’s study are repetitions, we might conclude that corrections by way of repetition of isolated learner errors are not preferred by FL teachers in general, whether their focus is on form or on forms.

Table 2 shows a summary of the findings presented in Table 1, in that it points out the different percentages of teacher initiated teacher corrections (CTs) compared to teacher initiated pupil corrections (CPs).

Within the present study where the focus is on forms (GFL) there are more CP moves (56%) than CT moves (44%). In Lyster and Ranta’s (1997) study two thirds of the correction moves turned out to be CTs (62%), because of the rather high percentage of recasts (55%). These findings

Table 2. The distribution of the CTs and CPs as % of the total number of feedback moves

	GFL (Flanders)	Lyster & Ranta (1997)
CTs:	44 %	62 %
CPs:	56 %	38 %
	(n=394)	(n=687)

seem to indicate that in the focus on forms classrooms students had more opportunities for self- or peer-repair than in highly communicative experiential classrooms, where teachers more often seem to provide the students with the correct forms in contingent recasts. Doughty's (2001: 256) argument that negotiation of form could be considered a "focus on forms approach" seems to be confirmed by these findings.

The question is whether these are valid comparisons, since the two FLT contexts discussed differ almost in every way. Lyster and Ranta's (1997) study is based on audio-recordings of a variety of lessons in four Grade 4 French immersion classrooms in Canada. This means that the learners in their study were 4 to 5 years younger than the learners in the GFL study. Furthermore, apart from the differences in L1 and target language, we do not have more information on other factors such as the differences in the learners' proficiency levels, in classroom culture or teachers' attitudes with regard to oral error treatment. However, because the aim of Lyster and Ranta's study was to describe how teachers and students engage in error treatment during communicative interaction, formal grammar lessons were excluded from their investigation. During the lessons, all on text comprehension, no explicit grammar teaching took place. We might therefore conclude that the negative feedback on formal error in these classrooms can be interpreted as focus on form activities. In the GFL study, on the contrary, all teaching activities can be described as being focus on forms, the text comprehension activities included. It is this fundamental difference we are interested in with regard to "noticing the feedback".

*Table 3.* Uptake following teacher feedback as % of feedback type

	repair	needs-repair	no uptake
CT:			
1. Explicit corrections	26	22	52
2. Recasts	35	12.5	52.5
CP:			
3. Clarification requests	0	100	0
4. Metalinguistic feedback	46.8	51.2	2
5. Elicitations	47	51	2
6. Repetitions	0	0	0

The second step in our inquiry was taking a closer look at the effectiveness of the correction types leading to learner uptake. Table 3 refers to the patterns of learner uptake following the different feedback types. When there is no learner uptake, there is either topic continuation or the teacher might use another correction move in order to have the learners repeat the correct answer or correct the errors themselves.

Some of the GFL findings in Table 3 are similar to Lyster and Ranta's findings. There is a high frequency of no learner uptake following the recast as a feedback type (52.5% vs. Lyster and Ranta: 69%). Similarly, explicit corrections lead to uptake only 50% percent of the time. While in the immersion study the explicit corrections are twice as likely to lead to repair, the repairs and needs-repairs in the GFL study are more evenly distributed.

Another interesting finding is that metalinguistic feedback and elicitations are the most successful techniques for eliciting learner uptake: metalinguistic feedback and elicitations lead to learner uptake 98% of the time. In Lyster and Ranta's study all the elicitations (100%) lead to uptake and only 14% of the correction moves involving metalinguistic feedback lead to no uptake, which in comparison to the amount of no uptake following explicit corrections and recasts is a very low percentage.

The same holds for the percentage of no uptake following clarification requests in Lyster and Ranta's study (12%). In the GFL study all the clarification requests lead to uptake, but this finding should be dealt with carefully, since only 1.8% of all correction moves in the GFL study were clarification requests. In the immersion study the repetitions proved to be rather successful in eliciting uptake (78% of the time), but this result cannot be compared to the GFL study, in which no such repetitions were found. Finally, there is a fifty percent chance of the uptake resulting in repairs or needs-repairs following correction moves with metalinguistic feedback or elicitations.

The last and most important issue involves the question of whether all repairs can be considered equally effective indicators that learners have noticed the feedback. When learners correct or at least try to correct themselves or peers, we might conclude that they have actually noticed the feedback. However, according to Lyster and Ranta a repair which is simply a repetition of what the teacher has said does not necessarily imply that the feedback has been noticed or understood as such. A further breakdown of the data, based on their study, involves a separation of peer- and self repairs, on the one hand from repetitions and incorporations on the other. In Table 4 the peer- and self-repairs are referred to as "student-generated repairs".

Table 4. Student-generated repairs as % of feedback type

	<i>GFL (Flanders)</i>	<i>Lyster &amp; Ranta (1997)</i>
CTs:		
1. Explicit corrections	0	0
2. Recasts	0	0
CP:		
3. Clarification requests	0	28
4. Metalinguistic feedback	46.8	45
5. Elicitations	47	46
6. Repetitions	–	31

Table 4 shows the percentage of each feedback type leading to student-generated repairs. Compared to Table 3 the percentages of repairs do not change for the CP-feedback types, i.e. the clarification requests, the metalinguistic feedback the elicitations and the repetitions. These types cannot elicit learner repetition, because the correct form is not provided by the teacher. However, since both the explicit correction and the recast techniques provide the learner with the correct forms, they cannot lead to student-generated repairs and therefore the percentages in Table 4 are 0% respectively.

We observe that nothing can be said about the zero percentage of student-generated repairs following the clarification requests, because these appear so infrequently in the data. The percentages involving the student-generated repairs following the repetitions cannot be discussed either, because no such repetitions were found in the GFL study. Surprisingly similar are the percentages of student-generated-repairs following correction moves with metalinguistic feedback (46.8% for GFL vs. 45% for immersion) and elicitations (47% for GFL vs. 46% for immersion) in both studies.

## 6. Discussion and conclusion

The first research question with regard to the present study, i.e. what kinds of oral corrective feedback constitute a FLT context where the focus is on forms, can be answered as follows: teachers are more likely to push students

to correct themselves by means of elicitations and metalinguistic feedback. By contrast, in the experiential setting the teachers frequently correct the student errors themselves by means of recasts. Clarification requests and repetitions, correction moves often related to the process of negotiation of meaning, hardly ever occur in the analytic FLT context (the GFL study), which may already indicate a lack of meaningful interaction.

It could be argued that the negotiations of form discovered in the GFL classes are more likely to occur in an environment of third language instruction (GFL in Belgium) than in a setting of second language instruction (French immersion in Canada). Note that the learners in the GFL study were four to five years older than the learners in the immersion study. The age difference may not be the only reason why the GFL learners are more experienced in “negotiating the form” or in “talking about language”; they also have been learning two other foreign languages in a classroom environment (French and English) since they were 12 years of age. Negotiations of form may therefore function better with more experienced FL learners (Lochtman 2002a).

The answer to the second question, i.e. how can the different kinds of oral corrective feedback be accounted for in terms of noticing the feedback, shows us the following picture. The uptake-findings in the present study are very similar to the findings of Lyster and Ranta (1997). In  $\pm 52\%$  of the CTs there is no uptake at all, whereas with the CP moves more than 95% results in uptake and therefore in negotiation of form. The correction types appear to function similarly in both FLT contexts. According to Lyster and Ranta (1997) and Swain (1998) this would imply that in the lessons where the focus is on forms more opportunities for foreign language learning were created, because teachers make more use of CPs. But if we look at the amount of successful uptake, we notice that CPs result in successful uptake only in 50% of the time. This might indicate that the learners did not always interpret the CPs as correction moves (see also Chaudron 1986; Kleppin and Königs 1991; Havranek 1996, 1997, 1999). It appears that the correction moves consisting of recasts result in less unsuccessful uptake and may therefore lead to learners’ noticing the gap. Furthermore, some studies suggest that the absence of uptake after recasts does not necessarily mean that the “gap” was not noticed (see e.g. Mackey and Philp 1998; Ohta 2000). Mackey and Philp (1998), for example, found that the L2 learners from their experiment whose erroneous utterances had been corrected by means of recasts and who had not repeated the correct utterance in any kind of uptake attained the same positive results after the experiment as the learners who did use uptake. They argue that uptake is not a prerequisite for noticing the feedback.

The findings above leave us with the following dilemmas. First, Lyster and Ranta (1997) suggest that negotiations of form foster opportunities for classroom foreign language learning, because learners are likely to notice that they are being corrected and are pushed to produce output as well. The findings indicate that if the use of negotiation of form is encouraged in FLT, this may constitute a move towards the analytic end of Stern's (1990) continuum (see also Doughty 2001). Negotiation of form does lead to uptake as opposed to teacher generated corrections, however. The question then is whether a higher frequency of negotiations of form in experiential FLT, as suggested by Lyster and Ranta, is feasible without the instructional setting becoming an focus on forms one.

Second, the outcome of the negotiations was not very successful. More than half of the uptake contained a new error or still needed repair in some other way. This shows that the language learners only noticed the gap in less than 50% of the teacher initiated student-generated attempts at repair. If that is the case there might be a risk for confusion on the part of the language learner. However, because the amount of uptake is very high, we may at least conclude that the students have noticed that they were supposed to produce new (comprehensible) output.

Third, although the amount of uptake following recasts is much smaller than the amount of uptake following the negotiations of form, it is mostly correct. Recasts do not interrupt the flow of discourse and there can be a direct comparison between the learner's faulty utterance and the teacher's correct one. However, it is not possible to conclude from the fact that there is no uptake after recasts that students have not noticed the teacher's correction. It may be that they don't feel the need to repeat the recast.

Finally, in the GFL classrooms recasts were followed by repairs in 35% of the time, which was higher than was the case for the French immersion classrooms. This seems to confirm Nicolas, Lightbown and Spada's (2001: 720) hypothesis that "recasts appear to be more effective in contexts where it is clear to the learner that the recast is a reaction to the accuracy of the form". Since in the analytic FLCs the focus is on forms and the teacher often knows in advance what the student is going to say (for example as an answer to display questions), recasts are not always semantically related to the learners' utterance and therefore not part of semantically contingent speech, i.e. of expansions or extensions. Long (1996) therefore calls this kind of recasts "pseudo-recasts". Such recasts function more like explicit corrections, because in most cases both the teacher and the learners focused on an isolated target form for the explicit purpose of foreign language learning. In the same vein, the negotiations of form in analytic FLT

should be considered as “pseudo-negotiations of form”, because they also do not occur in meaningful interaction in analytic FLT. However, the discourse analytic method used in both studies cannot show the distinction between “real recasts” and “pseudo-recasts” on the one hand and “real” negotiations of form and “pseudo-negotiations of form” on the other hand. It should be investigated whether and how discourse analytic methods can be enriched in order to deal with this problem, or whether a totally different approach is needed which will allow us to compare the discourse in completely different instructional settings (see also Norris and Ortega 2000).

The question which remains to be answered is whether one kind of corrective feedback is preferable as a correction move as opposed to another. Further research will help us decide whether learner uptake can be considered an effective indicator that learners have noticed the feedback. As mentioned earlier, studies have suggested that the absence of uptake after recasts does not necessarily mean that the “gap” was not noticed and where the uptake embedded in the negotiations of form is concerned, in half of the instances, errors remain.

## References

- Cathcart, Ruth L. and Judy E.W.B. Olsen  
1976 Teachers' and students' preferences for correction of classroom conversation errors. In *On TESOL '75*, Marina K. Burt and Heidi C. Dulay (eds.), 41–53. Washington D.C.: TESOL.
- Chaudron, Craig  
1986 Teachers' priorities in correcting learners' errors in French immersion classes. In *Talking to learn. Conversation in second language acquisition*, Richard R. Day (ed.), 64–84. Rowley (MA)/London/Tokyo: Newbury House Publishers Inc.
- Doughty, Catherine  
2001 Cognitive underpinnings of focus on form. In *Cognition and second language instruction*, Peter Robinson (ed.), 206–257. Cambridge: Cambridge University Press.
- Doughty, Catherine and Elizabeth Varela  
1998 Communicative focus on form. In *Focus on form in classroom second language acquisition*, Catherine Doughty and Jessica Williams (eds.), 114–138. Cambridge: Cambridge University Press.
- Doughty, Catherine and Jessica Williams  
1998 Issues and terminology. In *Focus on form in classroom second language acquisition*, Catherine Doughty and Jessica Williams (eds.), 1–12. Cambridge: Cambridge University Press.

- Ellis, Rod  
 1994 *The study of second language acquisition*. Oxford: Oxford University Press.  
 1997 *SLA research and language teaching*. Oxford: Oxford University Press.
- Farrar, Michael J.  
 1992 Negative evidence and grammatical morpheme acquisition. *Developmental Psychology* 28: 90–98.
- Gordon, Peter  
 1990 Learnability and feedback. *Developmental Psychology* 26(2): 217–220.
- Havranek, Gertraud  
 1996 What happens to the feedback? In *Approaches to second language acquisition*, Kari Sajavaara and Courtney Fairweather (eds.), 147–156. Jyväskylä: University of Jyväskylä.  
 1997 Corrective feedback revisited. In Lourdes Díaz and Carmen Pérez (eds.), *Views on the acquisition and use of a second language. Eurosla'7: Proceedings of the 7th EUROSLA conference Barcelona*, 417–426. Barcelona: Universitat Pompeu Fabra.  
 1999 The effectiveness of corrective feedback: preliminary results of an empirical study. In *Aile: Proceedings of 8th EUROSLA conference Paris*, 189–206, Volume 2.
- Henrici, Gert and Brigitte Herlemann  
 1986 *Mündliche Korrekturen im Fremdsprachenunterricht*. (Materialien zur Lehrerfortbildung), München: Goethe-Institut.
- Kleppin, Karin and Frank G. Königs  
 1991 *Der Korrektur auf der Spur. Untersuchungen zum mündlichen Korrekturverhalten von Fremdsprachenlehrern*. Bochum: Universitätsverlag Dr. N. Brockmeyer.
- Krashen, Stephen D.  
 1985 *The input hypothesis.: issues and implications*. London: Longman.
- Lier, van Leo  
 1988 *The classroom and the language learner 'ethnography and second language classroom research*. London: Longman.
- Lightbown, Patsy M. and Nina Spada  
 1990 Focus-on-form and corrective feedback in communicative language teaching. Effects on second language learning. *Studies in Second Language Acquisition* 12: 429–448.
- Lightbown, Patsy M.  
 2001 Input filters in second language acquisition. In *EUROSLA Yearbook 1*, Susan Foster-Cohen and Anna Nizegorodcew (eds.), 79–97. Amsterdam/Philadelphia: John Benjamins.
- Lochtmann, Katja  
 1997 Kommunikativer Grammatikunterricht an den flämischen Gesamtschulen: realisierbare Norm oder Utopie. *Zeitschrift für Angewandte Linguistik* 27: 113–127.



- 2000      *Korrekturhandlungen im Fremdsprachenunterricht*. PhD. diss., Department of German Linguistics, Vrije Universiteit Brussel.
- 2002a      Negotiation in the foreign language classroom: How corrective feedback affects classroom interaction in a third language environment. In Marc Hooghiemstra (ed.), *Interactive CD-ROM – L3 Conference*. Leeuwarden: Fryske Akademy.
- 2002b      *Korrekturhandlungen im Fremdsprachenunterricht*. Bochum: AKS-Verlag.
- Long, Michael H.
- 1983      Does second language instruction make a difference? A review of research. *TESOL Quarterly* 17: 359–382.
- 1991      Focus on form: A design feature in language teaching methodology. In *Foreign language research in cross-cultural perspective*, Kees de Bot, Daniel Coste, Ralph B. Ginsberg and Claire Kramsch (eds.), 39–52. Amsterdam: Benjamins.
- 1996      The role of the linguistic environment in second language acquisition. In *Handbook of research on second language acquisition*, William C. Ritchie and Tej K. Bhatia (eds.), 413–468. New York: Academic Press.
- Long, Michael H., Shunji Inagaki and Lourdes Ortega
- 1998      The role of implicit negative feedback in SLA: Models and recasts in Japanese and Spanish. *The Modern Language Journal* 82: 357–371.
- Lyster, Roy
- 1994      La négociation de la forme: stratégie analytique en classe d’immersion. *The Canadian Modern Language Review* 47: 159–176.
- 1998a      Recasts, repetition and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition* 20: 51–81.
- 1998b      Negotiation of form, recasts and explicit correction in relation to error types and learner repair in immersion classrooms. *Language Learning* 48: 183–218.
- Lyster, Roy and Leila Ranta
- 1997      Corrective feedback and learner uptake: Negotiation of form in communicative classrooms, *Studies in Second Language Acquisition* 19: 37–66.
- Mackey, Alison and Jenefer Philp
- 1998      Conversational interaction and second language development: Recasts, responses and red herrings? *The Modern Language Journal* 82: 338–356.
- Nicolas, Howard, Patsy M. Lightbown and Nina Spada
- 2001      Recasts as feedback to language learners. *Language Learning* 51 (4): 719–758.
- Norris, John M. and Lourdes Ortega
- 2000      Effectiveness of L2 instruction: a research synthesis and quantitative meta-analysis. *Language Learning* 50 (3): 417–528.

Nystrom, Nancy Johnson

- 1983 Teacher-student interaction in bilingual classrooms: Four approaches to error feedback. In *Classroom oriented research in second language acquisition*, Herbert W. Seliger and Michael H. Long (eds.), 169–189. Rowley (Ma)/ London/Tokyo: Newbury House Publishers Inc.

Ohta, Amy

- 2000 Rethinking recasts: A learner-centred examination of corrective feedback in the Japanese classroom. In *The construction of second and foreign language learning through classroom interaction*, Joan K. Hall and Lorrie Stoops Verplaetse (eds.), 47–71. Mahwah, NJ: Erlbaum.

Saxton, Matthew

- 1997 The contrast theory of negative input. *Journal of Child Language* 24: 139–161.

Schmidt, Richard

- 1995 Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In *Attention and awareness in foreign language learning*, Richard Schmidt (ed.), 1–64. Honolulu: National Foreign Language Resource Center.

Schmidt, Richard W. and Sylvia N. Frota

- 1986 Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. In *Talking to learn. Conversation in second language acquisition*, Richard R. Day (ed.), 237–326. Rowley (Ma)/London/Tokyo: Newbury House Publishers Inc.

Stern, H. H.

- 1990 Analysis and experience as variables in second language pedagogy. In *The development of second language proficiency*, Birgit Harley, Patrick Allen, Jim Cummins and Merrill Swain (eds.), 93–109. New York: Cambridge University Press.
- 1992 *Issues and options in language teaching*. Oxford: Oxford University Press.

Swain, Merrill

- 1985 Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In *Input in second language acquisition*, Susan M. Gass and Carolyn G. Madden (eds.), 235–253. London: Newbury House.
- 1995 Three functions of output in second language learning. In *Principles and practice in applied linguistics: Studies in honour of H.G. Widdowson*, Guy Cook and Barbara Seidlhofer (eds.), 125–144. Oxford: Oxford University Press.
- 1998 Focus on form through conscious reflection. In *Focus on form in classroom second language acquisition*, Catherine Doughty and Jessica Williams (eds.), 64–81. Cambridge: Cambridge University Press.

# **Noticing and the role of interaction in promoting language learning**

*Folkert Kuiken and Ineke Vedder*

In this chapter the importance of qualitative research with regard to the investigation of the role of noticing and interaction in promoting second language learning is emphasized. The results are reported of an experiment on the acquisition of the passive form, carried out among 34 Dutch high school students in their fifth year of English. The experimental group (n=20) was given two dictogloss tasks, which consisted of the reconstruction in small groups of two texts read by the teacher (+ interaction). The control group (n=14) performed the same tasks, but this time the students had to reconstruct the texts individually (-interaction). Knowledge of the passive was established by means of a pre-test. After the treatment a post-test and a delayed post-test were administered. By means of a quantitative analysis it could not be demonstrated that recognition and frequency of use of the passive differ depending on the degree to which learners are encouraged to interact with each other. A qualitative analysis of the reflections and discussions between the learners makes clear that numerous instances of interaction lead to the noticing of passive forms.

## **1. Introduction**

Does focusing on a particular language form promote the acquisition of that form? This question has been the matter of debate of much recent research on the effects of instruction on L2 acquisition (e.g. Doughty and Williams 1998; Norris and Ortega 2000; Skehan and Foster 2001). A crucial role in these studies is played by the concept of noticing. As stated by Robinson (1995, 1996), Schmidt (1990, 1995) and Skehan (1998), there seems to be a connection between learners' noticing of linguistic forms in the input and successful learning. Another important role is attributed to interaction and collaborative dialogue, that is dialogue in which learners are engaged in negotiating meaning. In order to be understood learners have to clarify, explain and correct themselves and this is how they learn about constructing

a sentence (Swain and Lapkin 2000, 2001). Also metacognition is thought to have a facilitative effect on L2 acquisition. Metalinguistic reflections about language may help learners to better understand the relation between meaning, form and function (Long and Robinson 1998).

Some authors have presented considerable evidence in favour of the effect of noticing (Doughty 2001; Doughty and Williams 1998; Norris and Ortega 2000). Others suggest that this might have to do with the fact that most research on noticing has been couched in the context of a 'Focus on Form' approach. For this reason they stress the need for studies in which the role of noticing is investigated in a 'Focus on Forms' context (Sheen 1994; Trenkic and Sharwood Smith 2001). All in all, if one has to characterize the empirical results regarding the relationship between noticing, interaction, metacognition and interlanguage development, the word 'mixed' seems to be the most appropriate, as pointed out by Skehan & Foster (2001).

In this chapter we report on an experiment in which the effect of noticing and interaction in a Focus on Form context was examined. The study was based on Skehan's information processing model (Skehan 1998). The model emphasizes input processing and the effects of input features, via noticing, on the interlanguage system of the L2 learner. Noticing, interaction and metacognition are also the pillars on which the dictogloss procedure is based, a pedagogic technique in which L2 learners have to reconstruct a text that has been read to them. For that reason the participants in our study carried out a dictogloss task. The question we wanted to answer was whether the outcome of this task differed, depending on the degree to which learners were encouraged to interact with each other. The linguistic feature we focused on was the passive.

In an earlier study we had not been able to establish an effect of interaction (Kuiken and Vedder 2002a). In this study participants with various target languages were involved: students of Dutch as a second language coming from various linguistic backgrounds, Dutch high school students learning English and Dutch university students studying Italian. The design of this study allowed us to distinguish between groups using different strategies in the reconstruction of the original text; in retrospect, however, the search for an effect of interaction was complicated by the selection of three rather diverse groups acquiring different target languages. Another limitation of the study concerned the fact that the data were only analysed quantitatively. For the present study we selected a more homogeneous group, that is Dutch adolescents studying English at high school. The data were analysed not only quantitatively, but also qualitatively.

The focus of this chapter will be on the qualitative analysis. In Kuiken and Vedder (2002b) the quantitative results were presented. In order to be able to interpret the qualitative data better, we will give an overview of the quantitative findings. The chapter concludes with a discussion of the results and with a plea for more qualitative research on the effect of interaction in focus on form tasks. We will start by clarifying some of the basic issues involved in our study, that is noticing, interaction, metacognition, the dicto-gloss procedure and Skehan's information processing model.

## **2. Theoretical assumptions**

### *Noticing*

Lately much SLA research has investigated whether noticing a particular linguistic form may promote acquisition of that form (Schmidt 1990, 1995; VanPatten 1996, 2000; VanPatten and Cadierno 1993). For a definition of noticing we refer to Schmidt (1990), who describes it as 'conscious attention to input'. Noticing a linguistic form in the input is thought to operate as a necessary, though not a sufficient condition for processing and acquisition to take place. Swain (1998) points out that there may be several levels of noticing: learners may simply notice a structure in the target language due to the frequency or salience of that form (Gass 1988) or they may notice that the target language form is different from their own interlanguage (also called 'noticing the gap', cf. Schmidt and Frota 1986). Although Schmidt (2001) prefers to use the term noticing in a restricted sense, as a technical term equivalent to 'apperception' (Gass 1988) or to 'detection within selective attention' (Tomlin and Villa 1994) and proposes to separate it from metalinguistic awareness, he agrees with Leow (1997) that it may be useful to distinguish two forms of noticing: simple noticing (i.e. registration with awareness, indicated by a report or repetition) and noticing with metalinguistic awareness (i.e. cases in which a structure is noticed and put into question or discussed). Leow (1997) demonstrated that those who showed higher levels of awareness learned more than those who merely attended and noticed. Later we will come back to this distinction.

### *Interaction*

According to Swain's output hypothesis (1985, 1998), output may influence noticing and promote L2 acquisition. More precisely, learners may use their output as a way of trying out new language forms and structures. They may

use output just to see what works and what does not (Swain 1998). In other words: language production may prompt learners to deepen their awareness of grammatical and lexical rules. This may trigger cognitive processes that generate new linguistic knowledge and consolidate existing knowledge. An important function in this respect is attributed to interaction. If no external source is available, there is no need for the learner to test and verify his hypotheses about the target language. In interaction, however, learners are able to obtain useful information for testing their hypotheses from others. And when they have been offered feedback, learners also modify and re-process their output (Swain 1993). For this reason Swain and Lapkin (2000, 2001) stress the role of collaborative dialogue. Through such dialogue, learners engage in co-constructing their L2 and in building knowledge about it. Whenever the term 'interaction' is used in this chapter, it should be interpreted in the sense of 'collaborative dialogue'.

### *Metacognition*

According to Ellis (2000), specific properties of a task may predispose or induce learners to engage in certain types of language use and mental processing that are beneficial for acquisition. Particular language production tasks, such as problem-solving activities, may encourage learners to talk about the linguistic problems they encounter. This metatalk, or verbalization of problems in contexts in which learners are engaged in meaningful interaction, may positively affect the acquisition of L2 knowledge, since these kinds of activities may lead to greater metacognitive awareness. Metalinguistic reflection about language may therefore help learners to understand the relation between meaning, form and function. It is language used for cognitive purposes (Long and Robinson 1998). Metatalk allows the researcher to observe learners' working hypotheses as they struggle toward solving their linguistic problems. It provides us a window into the language learning process because much of what is observed will be language learning in progress.

### *Dictogloss*

Noticing, interaction and metacognition as important conditions for second language learning are the theoretical assumptions underlying the dictogloss procedure, in which learners are encouraged to reflect on their own output. As described by Wajnryb (1990), a short text is read at normal speed to a group of L2 learners. This text, which is either a constructed or an authentic one, is intended to provide practice in the use of particular linguistic forms

or constructions. While the text is being read, learners take notes; they then work together in small groups to reconstruct the initial text from their shared resources. After the reconstruction phase the final version is compared with the original text, and then analysed and commented upon by the teacher. It is hypothesized that while learners interact with each other, their language ability improves, as far as their morpho-syntactic, lexical and pragma-rhetorical skills are concerned.

*Skehan's information processing model*

Our study is based on a modified version of Skehan's model on information processing. In Skehan's model the concept of noticing occupies a central position. According to Skehan (1998) various influences affect noticing, such as the frequency and salience of the input, classroom instruction, task demands on processing resources, individual differences between learners in processing ability, and readiness to pay attention to certain linguistic forms. In short the components of the model can be characterized as follows (see Figure 1):

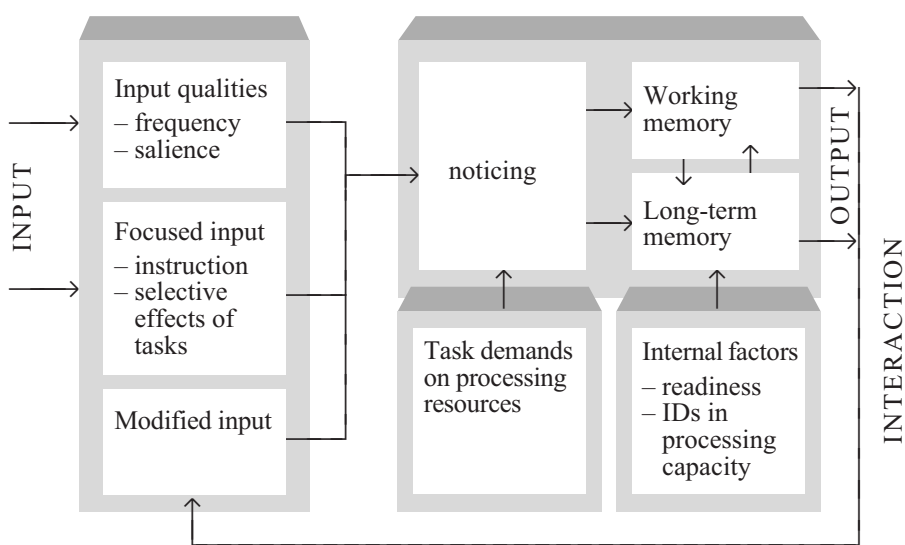


Figure 1. Types of noticing based on Skehan (1998)



- Input qualities: the more frequent and prominent a form, the more likely it is to be noticed in the input.
- Focused input: noticing may be influenced by instruction as learners may be prepared in such a way that less obvious aspects of the input become salient; similarly particular tasks may make certain language forms more salient.
- Task demands on processing resources: the cognitive complexity of a language task may influence noticing as well, in terms of making it more or less likely to occur.
- Internal factors: individual differences (IDs) between learners in processing ability, and readiness to pay attention to certain linguistic forms may also have an impact on noticing.
- Working memory and long-term memory: working memory is activated by the various influences operating upon noticing. The result of noticing then becomes available for modification and incorporation into long-term memory (Robinson 1995).
- Modified input: in line with Swain's output hypothesis we have added an extra component to Skehan's model, representing the effect that modified input may have on noticing. As stated above, noticing may be triggered by interaction, which may induce learners to search for alternative forms and to modify their output, in order to be understood by other learners (Long 1983, 1996). As a consequence, output becomes modified input and influences noticing. In Figure 1 this is indicated by the arrow that runs from 'output' via 'modified input' to 'noticing'.

Skehan's Information Processing Model applies to dictogloss in the following ways. The text read to the learners ('input' in Skehan's model) contains particular linguistic structures, which occur frequently and stand out prominently in the input. In the classroom learners' attention has already been drawn to them. Through its characteristics, the dictogloss task itself makes these structures more prominent ('focused input'). While interacting with each other, learners modify their output in order to be understood ('modified input'). In terms of cognitive complexity, the task demands of dictogloss on processing resources are probably neither too high nor too low, so it may be assumed that they will not overload the learners' internal system in such a way that noticing is less likely to occur (Robinson 2001). We will come back to this point in the discussion section. Between the learners of each group there will be differences in processing capacity and readiness to focus



on certain linguistic forms. During the text reconstruction phase, output and interaction may lead to an increased metacognitive awareness: learners find out what they know, should know and do not know about the target language. Therefore, noticing and co-construction of knowledge may lead to a change in the underlying interlanguage system. Working memory is activated and as a result the structures may be stored in long-term memory. Acquisition of the grammatical structures under investigation may therefore be facilitated.

### **3. Research question and design**

Our research question was whether noticing, triggered by interaction, lead to modified input. More particularly we investigated whether the outcome of a dictogloss task with respect to the passive differed depending on the degree to which learners were encouraged to interact with each other. Besides this quantitative question, which will be discussed in section 4, we were interested in how noticing took place and how different proposals for text reconstruction could be characterized (section 5).

The participants were Dutch high school students, aged between 16 and 18, who were in their fifth year of English. Throughout their school time the language education they received was based on principles of communicative language teaching, using a communicative course book. In their third and fourth year they had also studied a basic grammar of English, containing examples of the same three types of passive constructions they were confronted with during the experiment: type I passives are verbal constructions with one auxiliary (e.g. *were drawn* and *were made*), type II passives consist of two auxiliaries (e.g. *can be seen* and *could be made*) and type III passives of three auxiliaries (e.g. *must have been modelled* and *may have been created*). The students were distributed over two classes and were randomly assigned to an experimental group and a control group. Pre-existing knowledge of the passive in English was established by means of a detection test. This pre-test consisted of 32 sentences containing passive structures which the students had to underline. The test included eight passive structures of each type. One week after the pre-test the experimental group was asked to perform two dictogloss tasks. After the text was read out by the teacher, the learners had to reconstruct the original text in small groups consisting of three or four students. In total there were six of these groups. Those in the control group were given the same tasks, but they had to reconstruct the text individually, so that there was no possibility for

them to interact. Two texts were used: ‘The Stolen Painting’ (text A), a narrative text with events presented in a chronological order and ‘The Nazca Lines’ (text B), an expository text. Appendix 1 contains the original version of both texts, as well as an example of a reconstructed version of text A and B made by one of the groups, including their errors. The discussions between the students during the reconstruction phase were tape-recorded and transcribed. The whole procedure for the two dictogloss tasks took about one and a half hour. Immediately after the dictogloss tasks, a post-test similar to the pre-test was administered (see Appendix 2). Because not all the students who had done the pre-test were present in order to perform the dictogloss tasks and the post-test we ended up with an experimental group consisting of 20 students and a control group of 14. Two weeks later the learners took a delayed post-test (see Figure 2).

	Week 1	Week 2	Week 4
Experimental group (n = 20)	Pre-test	2 Dictogloss tasks (+ interaction) Post-test	Delayed post-test
Control group (n = 14)	Pre-test	2 Dictogloss tasks (– interaction) Post-test	Delayed post-test

Figure 2. Design of the study.

We hypothesized that the opportunity for the participants to interact with each other in the experimental condition (+ interaction) would result in a higher score on the post-test and delayed post-test and in the more frequent use of passive forms in the reconstructed texts than for the learners in the control group (– interaction). By means of a qualitative analysis we tried to determine how noticing of the passive took place and how different proposals for text reconstruction could be characterized.

#### 4. Quantitative analysis

Before answering the question whether the experimental group performed better than the control group on the post-tests and dictogloss tasks we performed some preliminary analysis taking into account the data of all the participants. We started by calculating the reliability scores of the tests by means of Cronbach’s alpha. These ranged from 0.7920 to 0.8711 and show that the data are reliable (see Table 1).

*Table 1.* Reliability analysis

	Pre-test	Post-test	Del. post-test
Cronbach's alpha	0.8525	0.7920	0.8711

As may be inferred from the mean overall scores on the three tests, the participants performed slightly better with time: on the pre-test about 55 items were scored correctly, on the post-test about 56 and on the delayed post-test about 58 (see Table 2).

*Table 2.* Mean overall test scores

Test	N	Minimum	Maximum	Mean	SD
Pre-test	34	44	69	55.7	7.1
Post-test	34	46	68	56.2	6.0
Del. post-test	29	44	69	58.4	7.0

It may not be surprising that in all the tests recognition of passive structures of type I and II was better than that of type III (see Table 3). But with time learners were better at identifying type III structures: their mean scores were 4.9 on the pre-test, 5.4 on the post-test and 5.7 on the delayed post-test.

*Table 3.* Mean overall test scores on type I, II and III structures

Type	N	Pre-test		Post-test		Del. post-test	
		Mean	SD	Mean	SD	Mean	SD
I	34	6.5	1.3	7.1	1.4	6.7	1.6
II	34	6.2	1.7	6.4	1.4	6.3	1.9
III	29	4.9	2.4	5.4	2.2	5.7	2.5

Although the participants had been randomly assigned to one of the two groups, there was a difference between the mean scores of the experimental group and the control group at the start (see Table 4). This difference still existed at the time of the post-test (57.9 for the experimental group versus

53.7 for the control group). With a score of about 58 for both groups, the difference between them had more or less disappeared at the time of the delayed post-test.

Table 4. Mean test scores for the experimental group and the control group

Group	N	Pre-test		N	Post-test		Del. post-test		
		Mean	SD		Mean	SD	N	Mean	SD
Experimental	20	57.6	7.7	20	57.9	6.5	18	58.5	7.2
Control	14	52.9	5.43	14	53.7	4.2	11	58.2	7.1

In order to test the effect of group interaction, an analysis of covariance (Ancova) was carried out. This made it possible to control for differences between the participants at the beginning of the experiment. No significant differences could be determined between the experimental and the control group at the post-test, as shown in Table 5 ( $p = .452$ ), nor at the delayed post-test, as can be seen in Table 6 ( $p = .407$ ). As a consequence, these findings do not support the hypothesis that giving learners the opportunity to interact with each other during a dictogloss task will result in a better score on the post-test or on the delayed post-test.

Table 5. Ancova for post-test with pre-test as co-variate

Source	Sum of squares	df	Mean square	F	Sig.
Pre-test	656.548	1	656.548	52.611	.000
Exper/Control	7.244	1	7.244	.580	.452
Error	386.859	31	12.479		

Table 6. Ancova for delayed post-test with pre-test as co-variate

Source	Sum of squares	df	Mean square	F	Sig.
Pre-test	484.488	1	484.488	14.096	.001
Exper/Control	24.455	1	24.455	.711	.407
Error	893.649	26	34.371		

A second analysis was performed in order to examine whether the two groups differed from each other in their use of passive structures in the reconstructed texts. Therefore the passive ratio was calculated. This ratio was determined by dividing the number of passive forms used in the reconstructed texts by the total number of verbal phrases used in the texts. For example, in the reconstructed text in Appendix 1, there are 5 passive forms out of a total of 12 verbal constructions, which gives a passive ratio of 0.4 (5 divided by 12). In order to determine the differences between the passive ratio of the experimental group and that of the control group in the reconstructed texts, analyses of covariance were carried out. These analyses did not reveal a significant difference between the two groups, for either text A ‘The Stolen Painting’ ( $p = .178$ ; see Table 7) or text B ‘The Nazca Lines’ ( $p = .244$ ; see Table 8). Thus, contrary to our expectation, the opportunity for interaction did not result in the more frequent use of passive sentences in the reconstructed texts.

*Table 7.* Ancova for passive ratio of text A ‘The Stolen Painting’ with pre-test as co-variate

Source	Sum of squares	df	Mean square	F	Sig.
Pre-test	1.406	1	1.406	.010	.920
Exper/Control	260.527	1	260.527	1.906	.178
Error	3964.211	29	136.697		

*Table 8.* Ancova for passive ratio of text B ‘The Nazca Lines’ with pre-test as co-variate

Source	Sum of squares	df	Mean square	F	Sig.
Pre-test	213.571	1	213.571	2.980	.094
Exper/Control	101.037	1	101.037	1.410	.244
Error	2221.907	31	71.674		

We then further analysed which types of structures were used in the reconstructed texts (see Table 9). It turned out that in the great majority of cases the simpler type I structures were used (on average 21.8 in text A and 24.9 in text B). Type II structures were used to a far lesser extent; there were some in text B (14.9), but hardly any in text A (3.2). The use of type III

structures was even more rare (1.7 in text A, 4.4 in text B). These findings confirm the results of our earlier study, in which we demonstrated that the morpho-syntax of the reconstructed sentences was generally less complex than the construction of the sentences in the original texts (Kuiken and Vedder 2002).

*Table 9.* Mean overall passive ratios obtained on text A ‘The Stolen Painting’ (n=32) and text B ‘The Nazca Lines’ (n=34)

Type	Text A ‘The Stolen Painting’		Text B ‘The Nazca Lines’	
	Mean	SD	Mean	SD
I	21.8	9.8	24.9	8.6
II	3.2	3.7	14.9	6.9
III	1.7	2.9	4.4	6.9
Total	26.6	11.7	44.1	8.6

As a consequence, we have to conclude that interaction during the reconstruction phase does not result in the better recognition of passives and a higher score in the detection test, nor in the more frequent use of these structures in the text reconstruction task.

## 5. Qualitative analysis

In the qualitative part of our research we analyzed the type of interaction, which was going on between the students. Our questions were: how does noticing take place, and how can different proposals for text reconstruction be characterized?

In the transcripts of the discussions we have identified all cases in which noticing of a passive structure was taking place. In total we found 79 of these cases. The discussions between the students were mainly in Dutch, partly in English. The examples 1-10 on which we will report below have been translated from Dutch. The relevant structures are in italics. Quotation marks indicate the utterances, which were spoken in English; formal errors have not been corrected.

Example (1) is a clear illustration of the way in which collaborative dialogue may take place and how learners are co-constructing L2-knowledge: by putting together several linguistic solutions proposed by each of them,

the students are building their sentences and reconstructing bit by bit the content of the text read to them by the teacher. In a joint effort they search for meaning and try to express this meaning in a correct form.

(1) Collaborative reconstruction of content

- Joanne: 'He... was... first... show... shown...'  
 Janis: '... the painting'.  
 Joanne: '... the money'.  
 Janis: Is that so? Oh yeah.  
 Joanne: 'And... ehm, after..., and... after...'  
 Jarnis: '... after that...'  
 Joanne: '... after that... they... went... into... a barn...'  
 Janis: '... to get the painting...'  
 Joanne: '... where...'  
 Janis: '... the painting was'.  
 Joanne: '... the painting... was shown...'  
 Janis: '... to the buyer'

Example (2) illustrates how different linguistic domains may be touched upon during one and the same discussion. Dirje wants to say: 'The Nazca Lines were made very precisely', but she doesn't know how to say 'precisely' in English. She receives help from one of her fellow students, who tells her that it is 'precise' in English. In Dutch, however, contrary to English, there is no formal distinction between the adjective and the adverb, so when Flora May realizes that in the sentence which Dirje is trying to formulate, 'precise' has the grammatical function of adverb, she adds *-ly* ('precisely'). Although Koen comes up with a synonym ('exactly'), Flora May and Dirje, ignoring his intervention, go on with the syntactic construction of the sentence. This results in the choice of 'were made'. So, in (2) the students first set off with a lexical discussion, then they focus on a grammatical and on a lexical issue, and then finally come back to grammar again.

(2) Shift from one linguistic domain to another

- Dirje: They eh, they made eh the, the... pictures  
 with a... lot of... LEXICON  
 I don't know how you say *precies* in English.  
 What is *precies*?  
 Flora May: 'Precise'?

Koen:	What?	
Dirje:	'Precise'? Okay.	
Flora May:	'-ly'.	GRAMMAR
Koen:	'Exactly'.	LEXICON
Flora May:	'Are made', isn't it? 'Very precisely'.	GRAMMAR
Dirje:	Yes, 'are made. Were made'.	

The examples (1) and (2) illustrate the following characteristics of collaborative dialogue:

- students help each other; if one student does not know how to continue, another comes up with a new proposal;
- there is a constant shift from one linguistic domain to another (in this case from lexicon to grammar and vice versa);
- a discussion about a particular problem often results in the use of an alternative form.

Let us consider now our first question: how does noticing take place? In order to investigate the role of noticing, we made a distinction between 'simple noticing' and 'elaborate noticing' (see section 2). By 'simple noticing' we refer to passives which are mentioned with more or less emphasis, but not discussed by the students. In case of 'elaborate noticing' passives are put into question, discussed and alternative structures may be proposed (see Leow 1997).

Examples (3) and (4) illustrate simple and elaborate noticing respectively. In (3) the passive 'were not discovered' is mentioned, but not discussed: this is an example of simple noticing.

### (3) Simple noticing

Janis:	'Yes, okay, the Nazca lines. Okay, the Nazca Lines <i>were not discovered</i> ... until... the nineteen-thirties'.
Joanne:	'Until... the nineteen-thirties'.

In (4) the students hesitate between 'the ancestors were given it', 'it was given by her ancestors' or 'it was given to her ancestors'. In the end they decide to go for the latter solution: an example of elaborate noticing. Note that in (4), just like in (2) meaning precedes form: first the participants search for meaning and decide which is the right content and then a decision about the right form is made. Again, the discussions about grammar and



lexicon are intertwined: Thomas doesn't know whether to use 'ancestors', or 'sisters', which sounds to him somewhat similar.

(4) Elaborate noticing

- Thomas: 'Her sisters', what, what, 'ancestors', what, what?  
 Orla: 'Ancestors...'  
 Thomas: '... were given it', shouldn't it be something like, eh, let's see, 'it was given by her ancestors' or so?  
 Björn: No, but how did these 'ancestors' get it, that's the point.  
 Orla: These 'ancestors' got it from the 'artist'.  
 Thomas: Oh.  
 Björn: And they, and they wanted to hand it over themselves...  
 Thomas: So, shall we say then, 'it was given to her ancestors'?  
 Björn: Yeah, okay.

Table 10 shows the number of cases of simple and elaborate noticing of passive structures for all the six groups (in total 79) during the reconstruction of 'The Stolen Painting' and 'The Nazca Lines'. As can be seen from the table, there were almost twice as many instances of elaborate noticing (51 versus 28). Moreover, in 'The Nazca Lines' the total number of cases of noticing, compared to 'The Stolen Painting', is nearly twice as high (50 versus 29).

Table 10. Number of cases of simple and elaborate noticing for all the groups (G1–G6)

Type of noticing	The Stolen Painting	The Nazca Lines	Total
Simple	12	16	28
Elaborate	17	34	51
Total	29	50	79

From this last finding we may infer that the process of noticing may be affected by factors like genre (narrative versus expository), and linguistic difficulty (for instance the amount of abstract versus concrete words). This has also been argued by Gass (1997, 2002), who investigated the different role that noticing played depending on the L2 proficiency level and the particular domain of language: morphology, syntax or lexicon. Noticing may also be influenced by pre-existing knowledge and familiarity with the

subject of the text: it could be hypothesized that an ‘unfamiliar’ subject might trigger more discussions between the participants and might therefore give way to more instances of noticing.

With regard to the number of cases of simple and elaborate noticing for each of the groups (G1–G6), there turn out to be considerable differences, as shown by Table 11: for group 1 and group 5 we find the highest number of instances of noticing (with totals of 20 and 16), for group 2 and group 4 the lowest (with a total of 9) each. These findings are in line with the results of our former study where we concluded that learners differed in the extent to which their attention was dedicated to language form (Kuiken and Vedder 2002). The differences are probably due to factors having to do with the composition of the groups. Students in the experimental condition were randomly assigned to one of the six groups, which may have resulted in differences in L2 proficiency level between the groups. Some students also seemed to have taken a more active part in the discussion than others. As demonstrated by Storch (2002) in a study on the effect of role relationships on language uptake in dyadic interaction, noticing may be influenced by factors having to do with status, social power, and personality. We will come back to this effect of group dynamics in the final section of the chapter.

*Table 11.* Number of cases of simple and elaborate noticing for each group (G1–G6)

Text	G1	G2	G3	G4	G5	G6
The Stolen Painting	5	6	2	3	8	5
The Nazca Lines	15	3	10	6	8	8
Total	20	9	12	9	16	13

Examples (5) and (6) illustrate how different linguistic proposals are commented on and discussed, which often results in a modification of the structure that had been originally proposed: an active construction may be substituted by a passive. In (5) the active construction proposed by Fatim (‘two men had sto..., had, yes’) is substituted by ‘had been stolen’. In spite of Fatim’s objection that it is not necessary to repeat exactly what the teacher said, Kirsten sticks to her proposal to use a passive. As a result of the discussion, ‘had stolen’ is replaced by ‘had been stolen’.

(5) Substitution of an active structure by a passive

- Fatim: Ehm, so 'two men had sto..., had', yes.  
Kirsten: No, 'tried to s..., to sell a painting', or something like that, 'that had been stolen', or so.  
Fatim: 'The... two men tried to sell a stolen, a stolen painting,' let's put it like this.  
Myrthe: Yes, but...  
Fatim: It's not necessary to repeat word by word what the teacher said, I mean...  
Myrthe: No, that would be impossible, but ehm..., let me think.  
Kirsten: I'm sure it has to be 'which had been stolen', or something like that.

However, also the opposite occurs, as demonstrated by example (6). In (6) the active construction 'could create' is changed into the passive 'can be constructed' and finally back again into the active structure 'could construct'.

(6) An active structure is substituted first by a passive and then again by an active structure.

- Lovella: Yes, so, 'that s..., that somebody *could create* this', hold on, ehm... 'it seems impossible to scientists of nowadays that somebody ehm... *could create* this pictures, ehm... with the science of three thousand years ago'.  
Fabe: Yeah.  
Lovella: Ready. But... ehm... '*can only be constructed... could construct... con... struct?*'  
Fabe: Also... Yeah.  
Lovella: Okay.

Let us turn now to the second research question, which was how different proposals for text construction can be characterized. Example (7) shows how the type I passive 'was given' is replaced by the more complex type II structure 'should be given'.

(7) A type I passive becomes a type II passive.

- Denise: 'And it *was given* to her as a wedding present... *should be given*...'  
Maarten: 'It *should be given* to her at her..., yes, at her wedding.'

In (8) the type I structure ‘they are made’/‘were made’/‘were created’ is changed into the type III structure ‘can/could have been made’. In this example the discussion between Maarten, Denise and Tin Choi regarding the way in which the Nazca Lines were made is integrated with a discussion about the passive and a discussion about syntax, in which explicit reference is made to rule knowledge of where to put the temporal constituent ‘three thousand years’: in first or final sentence position.

(8) A type I passive becomes a type III passive

Denise: Or... or were they made, by stones, oh...

Maarten: No, the... the... they were made by removing stones..., ‘*they must have... they...*’, how do you say that? Are they just supposed to be made in this way or... Do people know it for sure?

Denise: Yes, ‘*they...*’

Maarten: They are made... They are made 3000 years ago by removing huge..., by removing stones, yes. ‘*They were made...*’

Denise: ‘...*Created*’.

Maarten: Or ‘*created*’, yes... no... Probably time comes first... ‘Three thousand years’.

Denise: Yes, time first. No, I think it should be put at the end? Or right at the beginning? No, right at the end, it has to be right at the end of the sentence.

Tin Choi: Yes.

Maarten: Oh.

Denise: Yes. ‘By... stones?’ What did you say?

Maarten: ‘By digging away stones and rocks’, yes, ‘rocks’, I guess, ‘rocks and sand’.

Denise: ‘And...’

Maarten: ‘And stones and sand’.

Denise: ‘Sand’, yes.

Maarten: Ehm...

Denise: And ehm...

Maarten: Yes. ‘They are so huge... that they *can... could have been made... eh... only could have made*’. Yes.

In (9) the learners switch from the type I structure ‘were made’ to the type II passive ‘can’t be seen’, then try the active ‘to see’, after which they come

back again to the type II structure 'it can be seen'. What is interesting in this last example is that Orla first formulates a passive sentence in Dutch ('it can only be seen from a great height'), which is then translated by Thomas into English.

- (9) A type I passive becomes subsequently a type II passive, an active structure, and again a type II passive

Thomas: Ehm..., 'they were...' ehm... 'they *were made* about three hundred, three thousand years ago'. Ehm..., how could we say that from a great height they could not be... 'That...' No... ehm..., 'The peo...'

Björn: We can say that you can only see them from a great height.

Orla: 'They *can't be seen?*', no..., is that right?

Thomas 'The people eh, the people on the ground eh...'

Björn: There was something with, ehm, I don't know... Do you know it?

Thomas: 'It's impossible for people on the ground *to see* the pictures...'

Björn: 'On the ground?'

Thomas: [writing] 'It is... im-pos-si-ble... for people... on... the... ground...'

Orla: And now..., ehm... It can only be seen from a great height.

Thomas: Yes, '*it can... be seen*'.

In example (10) the participants are discussing what happened to the protagonist of 'The Stolen Painting': *was* he shown the painting or *had* he [himself] shown the painting [to other people]? As illustrated by this example, meaning is established, prior to form. After having agreed upon the meaning they want to express, students decide which linguistic form is the correct one: 'he was showed or he was shown?' Also in example (4), where the discussion developed more or less along the same lines, meaning precedes form: first the learners made clear that 'the ancestors got the painting from the artist', and then they decided that the construction 'it was given *by* her ancestors' was incorrect and should be replaced by 'it was given *to* her ancestors.' Another interesting point is that in (10), similarly to what happens in (9), the use of an L2 passive is based on an L1 structure. In this example the passive construction 'they are made by removing big stones,' initially formulated in the mother tongue, is translated by Maarten himself from Dutch into English.

## (10) Meaning precedes form

- Lovella: Let's say 'At first, ehm, he *was shown* the painting', because the... the painting was first, ehm...
- Fabe: 'He *had shown*', I think.
- Lovella: No, it should be 'he *was shown*', he was the one..., let's say, ehm, the painting was shown to him.
- Fabe: O yes, yeah.
- Lovella: '*Shown*...'
- Hester: But doesn't this mean that he himself was shown?
- Lovella: No, 'at first he *was shown*', he was...
- Fabe: Otherwise it should be 'he *was showed*'.
- Hester: Yes.
- Lovella: Yeah, 'he *was showed*', or no, 'he *showed*', yes, 'he *was showed*', yes, that's right.
- Hester: Okay.
- Lovella: 'At first he *was shown* the painting... in the banner', I just add that, who cares if it's wrong! [laughs]
- Fabe: 'Banner' is a kind of hangar I think or not?
- Lovella: Yes, but in any case '*was shown* the painting in a banner...'

To sum up: the qualitative analysis of the discussions shows that collaborative dialogue in many cases favours co-construction of L2 language building. There are numerous instances of interaction leading to elaborate noticing. This noticing leads in general to new linguistic proposals and to modified input, which often results in the use of more complex verbal forms (examples 4, 7, 8, 9). Even in cases where preference is given to a grammatically simpler form, like in (6), it is clear that noticing of passives has taken place. What all these examples demonstrate is that students are aware of the problems they encounter. Discussions on grammar are often mingled with discussions on other linguistic domains. Meaning generally precedes form: while searching for meaning students try to connect the correct form to the meaning they want to express (examples 4 and 10). In formulating an L2 passive or other syntactic constructions, students may rely on their mother tongue (examples 9 and 10), on implicit knowledge of L2 (example 5) or on explicit rule knowledge (example 8). For 'The Nazca Lines' there are much more examples of simple and elaborate noticing than for 'The Stolen Painting', which may have to do with differences in genre, linguistic difficulty and familiarity with the subject of the text. Finally, there are substantial differences between the six groups concerning the number of cases of noticing,

which are probably due to differences in language proficiency or to factors having to do with group dynamics.

## **6. Discussion**

Although by means of a quantitative analysis it could not be demonstrated that the opportunity for interaction during the reconstruction phase resulted in a better score on the detection test or in more frequent use of the passive in the reconstructed texts, the qualitative analysis has revealed that interaction about grammatical structures may stimulate noticing of a particular linguistic form. By putting into question their initial proposals students modify their output in order to find a better solution. In doing this they sometimes make reference to explicit knowledge of a particular rule, while in other cases they are relying on implicit knowledge or intuition. In some cases a more general searching strategy is used or learners rely on their mother tongue. In other cases avoidance behaviour can be observed, in so far as learners choose the construction they feel the safest about. Our data contain many examples demonstrating that students comment on each others' proposals. On some forms learners agree immediately, in other cases various alternative forms are put forward, discussed and sometimes put aside. There may be various reasons why some forms are rejected and others incorporated. It is possible, as hypothesized by Villamil and De Guerrero (1980), that some linguistic domains are more susceptible for incorporation of new forms than others. What might be investigated in future research is the question of which cases and why learners decide whether or not to accept new linguistic solutions. A more detailed analysis of the nature of the interaction which is taking place during the reconstruction phase of the dictogloss task should reveal how discussions between learners lead to more noticing and, in the end, to the use of more target structures. Studies like the one undertaken by Swain and Lapkin (2001) seem promising in this sense. Such a qualitative study should also make clear in which cases learners decide whether or not to accept alternative linguistic proposals and to incorporate them into their texts.

When we come back to our research questions, we may conclude that interaction triggers noticing. What has not been shown is whether this noticing leads in the end to acquisition. In this sense our study is in line with other studies that could be characterized as "mostly descriptive in nature, focusing primarily on occurrence per se rather than acquisition..." (Shehadeh 2002: 597). In terms of Skehan's Information Processing Model,

our study makes clear that the cycle noticing → output/interaction → modified input is completed several times during the discussion of one and the same problem. This is even true for those discussions that do not result in the use of an alternative or more complex form, when the outcome of a discussion is that learners decide to write down what has been suggested first.

In future research it might be useful to investigate more closely the various components of Skehan's model affecting noticing, in particular the influence of task demands ('cognitive complexity') on learners' processing resources. There are, however, competing accounts regarding the role of cognitive complexity. According to Robinson (2001) functional and structural complexity are associated with one another. In this view cognitively complex tasks would promote greater accuracy and greater complexity while simple tasks would lead to less complexity and lower accuracy. A contrasting viewpoint (Skehan 1998; Skehan and Foster 2001; VanPatten 1990) is that humans have limited information processing capacity and must therefore prioritise concern for content over concern for form. Tasks which are cognitively demanding in their content are likely to draw attentional resources away from language forms, encouraging learners to avoid more complicated linguistic structures in favour of simpler language forms. It would be interesting to explore these two contrasting views empirically in a follow-up study.

To explain the lack of a quantitative effect of interaction some limitations of the study have to be taken into account. First the number of participants involved in our study was relatively small and we had to deal with differences between the experimental group and the control group right from the beginning, as pointed out in section 4. Secondly, we have to consider the duration of the treatment (the two dictogloss tasks), which was about one and a half hour. It seems likely that language learners may need several opportunities to notice and work on a structure before it appears in their language production. Since the students had already been exposed on several occasions to the explanation and use of the passive in English, it may be assumed, though, that this knowledge was reactivated during the dictogloss tasks. It would nonetheless be interesting to see what happens when learners are given more dictogloss tasks over a longer period of time.

There seem to be various factors which may influence the process of noticing. As discussed in section 4 of this chapter, there turned out to be considerable differences concerning the number of instances of noticing between the six groups in the experimental condition. Firstly, language proficiency seems to play a role. Both in Kuiken and Vedder (2002a) and in the present study the number and kind of strategies used, and the instances of simple and elaborate noticing seemed to depend on the level of L2 profi-



ciency. In our former study we assumed that the more learners discussed grammatical and lexical issues, the better would be the quality of the text: more noticing may lead to a lexically and grammatically more complex text. This idea could, however, only partially be confirmed. A possible explanation for this finding might be that some learners do not need to talk so much about language form, simply because they know that what they are writing is correct.

This brings us to a second factor that may influence noticing, namely group dynamics. As stated by Long (1996), interaction seems to take place especially in case of an information gap, when learners with different levels of L2 proficiency question each other's linguistic proposals. For less proficient learners, this information gap is certainly an advantage: they are able to profit from the correct solutions proposed by more advanced learners. The opposite, however, may not always be the case: an incorrect structure proposed by a less proficient learner may be accepted by other learners simply because he or she has a more extrovert personality and more social prestige (Storch 2002). This complicates the comparison of the group results of the dictogloss text with the individual results of the pre-test, post-test and delayed post-test.

A third factor, which seems to affect noticing, is text difficulty. In our first study the number and kind of strategies used and the number of cases of simple and elaborate noticing seemed to be related to text difficulty, as these numbers differed within the same group from one text to another. Also in the present study there turned out to be an effect of text difficulty, genre and familiarity of subject. As shown by the different passive ratios in 'The Stolen Painting' and 'The Nazca Lines', results of similar dictogloss tasks may vary. As demonstrated in Table 9 for the Nazca lines, there were more examples of simple and elaborate noticing and also the number of passives actually used in the reconstructed version was much higher. So noticing is probably influenced by factors like genre, subject, and text difficulty.

We may conclude that, although the quantitative analysis does not show significant gains when learners are given the opportunity to interact, the results of the qualitative analysis seem to point to an effect for interaction. What our examples show is that in the dictogloss task, discussions focused on meaning often trigger discussions on form. While engaged in meaning negotiation learners are continuously reflecting on form. Together they learn something new, or they consolidate what they already know about L2. A qualitative analysis is, in our opinion, a necessary procedure to shed light on the process of noticing and collaborative learning, and to give an insight into the way in which form-meaning connections are established.

## Appendix 1

### Original and reconstructed texts

(In the texts below **type I** structures are underlined, **type II** structures are printed bold, while **type III** structures are underlined and printed bold.)

#### *Text A: 'The stolen painting' – Original text*

Two men tried to sell a painting that **had been stolen**. The painting was owned by Mary Jones, aged 84, who said it **may have been presented** to one of her ancestors by the artist himself. It **should have been given** to her in 1941 as a wedding present, but due to threat of war, she did not receive it until 1945. One of the accused men, Mr X, who **cannot be named** for legal reasons, pleaded guilty. He told the police that even though a high price **should have been paid** for the painting, he was willing to sell it cheap, in order to get rid of it. A meeting was arranged with a potential buyer, at an airfield near Liverpool. There the money for the painting **was to be flown in** and **exchanged**, but the airfield **had been staffed** by police officers in plain clothes. Mr X took the painting to the airfield and was shown the money in a suitcase. The buyer was then taken to see the painting in a barn, where he was arrested by the police.

Text adapted from D. Willis (1991), *Collins Cobuild Student's Grammar*. London: HarperCollins Publishers, p. 171.

#### *Text A: 'The stolen painting' – Reconstructed text*

Two men tried to steal a painting, a painting who was owned by Mary Jones, aged 84 years old. She said that maybe an ancestor showed the painting to the two men. She **should have received** the painting as a wedding present in 1941, but she didn't receive it until 1945. Mr X (whose name **can't be told** for legal reasons) who was accused of stealing it, wanted to sell the painting for a cheap price to get rid of it. At Liverpool airport the exchange took place. The buyer was flown in there with the money. At first, he was shown the painting in a banner, then he showed the money in a suitcase to Mr X. But the police had already surrounded the airport and they arrested Mr X.

*Text B: 'The Nazca lines' – Original text*

The Nazca Lines were not discovered until the 1930s, when they were first noticed by airplane pilots flying over Peru's Atacama Desert. They consist of huge pictures, several kilometers in size, that were drawn in the desert. According to their appearance they **must have been modelled** on birds, spiders, and abstract geometrical designs. These pictures were made more than 3000 years ago by removing stones and dirt over large areas. The amazing thing about the Nazca Lines is that none of these pictures **can be seen** by people on the ground. They are so huge that they **can only be seen** from a great height. The pictures **must have been constructed** with incredible precision. Exactly how such precise measurements **could be made** still **hasn't been** satisfactorily **explained**. It seems impossible that the primitive construction techniques that existed 3000 years ago **could have been used** to create such gigantic, perfectly constructed designs. Until now it is still unclear by whom these gigantic pictures **may have been created** and why.

Text adapted from: S.H. Tawlis (1993) *Grammar dimensions. Form, meaning and use. Book three*. Boston, MA: Heinle & Heinle.

*Text B: 'The Nazca lines' – Reconstructed text*

The Nazca lines were not discovered until the 1930's. They were discovered by airplane-pilots, when they were flying over Peru's Atacama Desert. They found huge pictures, which were modelled 3000 years ago. By the moving of stones, pictures of birds, spiders etc. **could be seen**. But the pictures **could only be seen** from a huge height. It seems impossible to the scientists of nowadays, that somebody could construct this pictures with the science of 3000 years ago. They still don't know the reason why they made those pictures and who made them.

## Appendix 2

### Post-test (instruction and sentences 1-10)

Underline the passive constructions in the following sentences.

Examples:

*A concert to raise money for the refugees will be held next Saturday night.*

*Several years ago Richard Anderson, a former astronaut, was informed by his superior at the aircraft corporation that he could no longer be a test pilot.*

1. According to a recent survey, out of every guilder a Dutchman spends on drinks, twenty-eight cents is spent in bars.
2. According to the president the reconstruction of the railway should have been finished by the end of last year.
3. Many engineers believe that wind energy will be used extensively in the next decade.
4. Registration forms should be returned by 8 July.
5. The front door was shut by our babysitter when we left the house.
6. Because the sun was shining so brightly, the blinds of the museum should have been closed.
7. Lady Di will never be forgotten. She was a woman whose name will go down in history.
8. The doors should have been painted while we were away.
9. The tents were all blown away during the storm by the heavy winds.
10. Nine of the pupils were unable to do the test. I think that the test should have been cancelled.

## References

- Doughty, C.  
2001 Cognitive underpinnings of focus on form. In *Cognition and second language instruction*, P. Robinson (ed.), 206–257. Cambridge: Cambridge University Press.
- Doughty, C. and J. Williams (eds.)  
1998 *Focus on form in classroom second language acquisition*. Cambridge: Cambridge University Press.
- Ellis, R.  
2000 Task-based research and language pedagogy. *Language Teaching Research*, 4 (3): 193–220.
- Gass, S.  
1988 Integrating research areas: A framework for second language studies. *Applied Linguistics*, 9: 198–217.  
1997 *Input, interaction and the second language learner*. Mahwah NJ: Erlbaum.  
2002 *The role of context and SLA*. Paper presented at Form-Meaning Connections in Second Language Acquisition. Chicago.
- Kuiken, F. and I. Vedder,  
2002a Collaborative writing in L2: The effect of group interaction on text quality. In *New directions for research in L2 writing*, G. Rijlaarsdam, M. L. Barbier and S. Ransdell (eds.), 168–187. Dordrecht: Kluwer Academic Press.  
2002b The effect of interaction in acquiring the grammar of a second language. *International Journal of Educational Research*, 77: 343–358.
- Leow, R.  
1997 Attention, awareness and foreign language behavior. *Language Learning*, 47: 467–505.
- Long, M. H.  
1983 Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics*, 4: 126–141.  
1996 The role of the linguistic environment in second language acquisition. In *Handbook of second language acquisition*, W. Ritchie and T. Batia (eds.), New York: Academic Press.
- Long, M. H. and P. Robinson  
1998 Focus on form: Theory, research, and practice. In *Focus on form in classroom second language acquisition*, C. Doughty and J. Williams (eds.), 15–41. Cambridge: Cambridge University Press.
- Norris, J. M. and L. Ortega  
2000 Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50 (3): 417–528.

- Robinson, P.  
 1995 Attention, memory and the noticing hypothesis. *Language Learning*, 45 (2): 283–331.  
 1996 *Task complexity and second language syllabus design: Data-based studies and speculations*. Special Issues of the University of Queensland Working Papers in Language and Linguistics.  
 2001 Task complexity, cognitive resources, and syllabus design: A triadic framework for examining task influences on SLA. In *Cognition and second language instruction*, P. Robinson (ed.), 287–318. Cambridge: Cambridge University Press.
- Schmidt, R.  
 1990 The role of consciousness in second language learning. *Applied Linguistics*, 129–158.  
 1995 Conscious and foreign language learning: A tutorial on the role of attention and awareness in learning. In *Attention and awareness in foreign language learning*, R. Schmidt (ed.), 1–64. Honolulu: Second Language Teaching & Curriculum Center, University of Hawai'i.  
 2001 Attention. In *Cognition and second language instruction*, P. Robinson (ed.), 3–32. Cambridge: Cambridge University Press.
- Schmidt, R. and S. Frota  
 1986 Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. In *Talking to learn: Conversation in second language acquisition*, R. Day (ed.), 237–322. Rowley, MA: Newbury House.
- Sheen, R.  
 1994 A critical analysis of the advocacy of the task-based syllabus. *TESOL Quarterly*, 28 (1): 127–151.
- Shehadeh, A.  
 2002 Comprehensible output, from occurrence to acquisition: An agenda for acquisitional research. *Language Learning*, 53 (3): 597–647.
- Skehan, P.  
 1998 *A cognitive approach to language learning*. Oxford: Oxford University Press.  
 2001 Tasks and language performance assessment. In *Researching pedagogic tasks. Second language learning, teaching and testing*, M. Bygate, P. Skehan and M. Swain (eds.), 167–185. Harlow: Pearson Education.
- Skehan, P. and P. Foster  
 2001 Cognition and tasks. In *Cognition and second language instruction*, P. Robinson (ed.), 183–205. Cambridge: Cambridge University Press.
- Storch, N.  
 2002 *Role relationships in dyadic interactions and their effect on language uptake*. Paper presented at Form-Meaning Connections in Second Language Acquisition. Chicago.

- Swain, M.  
1985 Communicative competence: some roles of comprehensible input and comprehensible output in its development. In *Input in second language acquisition*, S. Gass and C. Madden (eds.), 235–253. Rowley MA: Newbury House.  
1993 The output hypothesis: Just speaking and writing are not enough. *The Canadian Modern Language Review*, 50 (1): 158–164.  
1998 Focus on form through conscious reflection. In *Focus on form in classroom second language acquisition*, C. Doughty and J. Williams (eds.), 64–82. Cambridge: Cambridge University Press.
- Swain, M. and S. Lapkin  
2000 Task-based second language learning: The uses of the first language. *Language Teaching Research*, 4 (3): 251–274.  
2001 Focus on form through collaborative dialogue: Exploring task effects. In *Researching pedagogic tasks. Second language learning, teaching and testing*, M. Bygate, P. Skehan and M. Swain (eds.), 99–118. Harlow: Pearson Education.
- Tomlin, R. and V. Villa  
1994 Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition*, 16: 183–203.
- Trenkic, D. and M. Sharwood Smith  
2001 *Reevaluating theoretical and methodological aspects of Focus on Form research*. Paper presented at Eurosla 11, Paderborn.
- VanPatten, B.  
1996 *Input processing and grammar instruction*. New York: Ablex.  
2000 Thirty years of input (or intake, the neglected sibling). In *Social and cognitive factors in second language acquisition. Selected Proceedings of the 1999 Second Language Research Forum*, B. Swierzbin et al. (eds.), 287–325. Somerville: Cascadilla Press.
- VanPatten, B. and T. Cadierno  
1993 Explicit instruction and input processing. *Studies in Second Language Acquisition*, 15 (2): 225–243.
- Villamil, O. S. and M. C. M. D. Guerrero  
1998 Assessing the impact of peer revision on L2 writing. *Applied Linguistics*, 19 (4): 491–514.
- Wajnryb, R.  
1990 *Grammar dictation*. Oxford: Oxford University Press.





## **Interactional strategies for interlanguage communication: Do they provide evidence for attention to form?<sup>1</sup>**

*María del Pilar García Mayo*

Recent research has claimed that second language (L2) learners need access to positive input of L2 forms and their relation to meaning and to negative input about what is not in the L2 (Long 1996). Besides, it has also been claimed (Swain 1985, 1995) that learner production of modified output is also necessary for L2 mastery. Among the ways in which interaction can be modified, negotiation of meaning has been shown to be particularly productive. The present study considers if learners' needs in an English as a Foreign Language (EFL) setting can be addressed through interaction other than negotiation of meaning. The performance of seven dyads of advanced EFL learners and of seven dyads of those learners with English native speakers was analyzed on two types of communication tasks: information gap and decision making. Information gap tasks were used for their established effectiveness in providing learners with opportunities towards comprehension, feedback and interlanguage modification. Decision making tasks were used as a way to generate an exchange of ideas as the learners engage in opinion, argument and decision-oriented outcomes. Results of our research show that during their interaction advanced learners were resourceful in using strategies such as repair (self- and other-) and completion. They were able to draw from their interlanguage store both to correct and clarify message meaning and complete each other's messages. Both strategies allowed for further exchange of messages, moved the interaction along and seemed to facilitate the development of grammatical and lexical features. However, the numerous morphosyntactic imprecisions observed in the learners' output strongly suggest the need to devise form-focused tasks that target the specific grammatical and lexical features that still need to be mastered at this stage.

## 1. Introduction

The needs of L2 language learners have been analyzed and illustrated in numerous studies to date (see Gass and Selinker 1994; Pica 1994; Swain 1995; Long 1996; Mitchell and Myles 1998; Gass, Mackey and Pica 1998, among others). According to Long (1996), learners need access to (i) positive input of L2 forms and their relation to meaning and (ii) negative input about what is not in the L2. Adding to the learners' input needs, Swain (1985, 1995) has claimed that learner production of modified output is also necessary for L2 mastery. Learners need access to situations in which they can produce meaningful L2 output and modify it toward greater comprehensibility.

Among the ways in which interaction can be modified, negotiation of meaning has been shown to be particularly productive. Research has shown that when interaction is modified through the triggers, signals and responses of negotiation, conditions for L2 learning are enhanced considerably. Pica (1994) has reviewed in detail how negotiation can accomplish a great deal for second language acquisition (SLA). It can make input comprehensible to L2 learners (Pica, Young and Doughty 1987; Gass and Varonis 1994); help them modify their own output (Holliday 1995; Linnell, 1995) and provide opportunities for them to access L2 form and meaning. Negotiation of meaning plays an important role in SLA because (i) it triggers attentiveness and involvement, both of which are necessary for successful communication, and (ii) through negotiation learners are made aware of errors in their speech (Gass and Selinker 1994: 219). Negotiation is what makes learners aware of the mismatch between their output and the target language forms.

Previous research (García Mayo and Pica 2000a, 2000b) revealed that, overall, interaction between advanced learners engaged in pairwork on communicative tasks provides as much modified input, feedback and output as when interaction between learners and native speakers takes place. Learners could offer each other modified L2 input and grammatically accurate feedback and could produce modified output but these features were very low in frequency though as negotiation, the usual vehicle for their generation, was seldom used during the learners' interaction. What was observed was that learners were able to convey comprehensible messages as requested by the tasks used, so there was little need for negotiation on their part. The learners were generally comprehensible, but often not target-like. Several grammatical imprecisions (related to subject/verb agreement, third person singular subject/object pronouns and possessives, article over- and under-suppliance, subject pronoun omission, adverb misplacement and misuse of prepositions) and lexical imprecisions were observed. But these

interlanguage features did not appear to interfere with learners' comprehensibility.

Therefore, the present study was undertaken to address the question of whether the needs of learners in an EFL setting could be addressed through interaction other than negotiation of meaning. The following research questions were advanced:

- (i) Do learners engage in other kinds of interactional modifications?
- (ii) If so, do these modifications address learners' needs for positive and negative input and production of modified output?
- (iii) Do these modifications enable them to focus on L2 form-meaning relationships?
- (iv) Are the form-meaning relationships the ones that are crucial to these students' morphosyntactic development?

## **2. Methodology**

The study reported on in this paper was carried out in the Basque Country (Spain) with students with an advanced level of English who are studying the language not only as a foreign language but, for many, as a third language as they are already bilingual in Basque and Spanish. These learners are enrolled as second year students in the four-year English Philology degree awarded by the University of the Basque Country. They receive instruction exclusively in English in all their degree-related subjects (Phonetics, Morphology, History of the English Language, Syntax etc). The program they follow in their English language classes is communicative (Savignon 1991), that is, during classroom time different types of activities are used to bolster input and encourage L2 production. There is an emphasis on using authentic language and on successful communication.

The subjects, data collection procedures and database used for this study are the same that appear in García Mayo and Pica (2000a, 2000b). The relevant information is summarized in sections 2.1 and 2.2.

### **2.1. Subjects**

Subjects were fourteen advanced learners of English (seven males and seven females), with TOEFL scores in the 580–630 range, and seven female native speakers (NS) of English. The learners were assigned to one

of seven dyads of learner-learner (L-L) interactants. Six of the NS were North American college students from three different universities, all of them speakers of standard American English. The other NS was a British college student who had come to the Basque Country as an ERASMUS student.<sup>2</sup> They were assigned to one of the seven native speaker-learner (NS-L) dyads. Dyadic distribution of subjects was as follows: 2 were male L/male L pairs, 2 female L/female L pairs, 3 male L/female L pairs, 5 female NS/female L pairs and 2 female NS/male L pairs.<sup>3</sup>

The learners ranged in age from 19–33 (median 22). Their median of exposure to English was approximately 10 years. The NS ranged between 19–22 (median 20 years). Assignment into dyads was based primarily on the participant's availability, which was constrained by class schedules.

## 2.2. Data collection procedures

Recordings were made in a period of approximately one month when the members of the dyads were available. They took place in a laboratory setting at their university. The researcher introduced the members of the dyads to each other, reviewed instructions for taping, advised them to read instructions carefully and left them to work. Four tasks were used: two information gap and two decision making tasks. The information gap tasks were used for their established effectiveness in providing learners with opportunities to work toward comprehension, feedback and interlanguage modification. This is because they are required to exchange information in order to reach the goal of the task. The decision making tasks were used as a way to generate an exchange of ideas as the learners engage in opinion, argument, and decision oriented outcomes.

The information gap task used with the L-L dyads was "The unlucky man", from P. Ur *Discussions that Work* (1996: 63). Individual learners were given five different vignettes from a ten-scene story, which they were then told to arrange into a story by exchanging information about the vignettes held uniquely by them. They were not allowed to view each other's pictures or the original ten-scene story until they completed the task.

The decision making task for the L-L dyads, "The desert island", was taken from S.A. Sadow's *Idea Bank* (1982) and Duff (1986). The learners were told to imagine they were on a sinking ship. The instructions relayed that there were rubber boats available for their rescue. However, the boats could hold only a limited amount of supplies and people. A small island could be seen in the distance. If their boat made it to the island safely, they

would need things to help them survive and were told to choose three items from each of the groups. The two members of the dyad had to decide, and agree completely, on which items to take and which to leave behind.

The L-NS dyads participated in two communication tasks as well. The content of these tasks was modified as seven of the students in the L-L dyads were now in the L-NS dyads. Their information gap task, also a picture task, was based on Mathematical games by Martin Gardner (© by Scientific American, Inc. Reprinted by permission in P. Ur, 1996: 62). Their picture sequence consisted of seven drawings; each member of the dyad had three of those and were allowed to see the seventh, remaining drawing. The task required members of the dyad to describe the scenes they held, and uncover the story line behind them. In the pictures they saw a man that had to take a goat, a wolf and a cabbage in a small boat from one island to another. Specific instructions were given as to which two animal/vegetable combinations could be left together in one of the islands. The members of the dyad had to come up with a logical order for the different scenes and discover how the man managed to solve his transportation problem with animals and vegetables intact.

The decision making task given to the L-NS dyads was “Choosing candidates” (The Law scholarship, from P. Ur 1996: 72). In this task the members of the dyad were asked to choose one candidate to be awarded an annual Law Scholarship. The dyads were provided with profiles of five candidates who had all attained similar grades on their university entrance exam. There was detailed information about the relative merits of each candidate: their personal backgrounds, needs, tastes and characters.

All these tasks were taken from actual published materials and were open-ended as far as expectations about linguistic features. The primary motivation for their choice was that they resembled the kinds of communicative activities typically employed in tertiary and university EFL classrooms. A total of six hours of recording were transcribed and coded and an utterance count (Chaudron 1988; Crookes 1990) was carried out: a pause of more than one second was an indication of an utterance boundary in the database analyzed.

### **3. Interactional strategies for interlanguage communication**

Two interactional strategies that might be of possible relevance to our first research question (Do learners engage in other kinds of interactional modifications?) were found in the database presented above, namely, learner

self-repair and completion strategies as a context of positive input. We examine each of them in turn in the following sections.

### 3.1. Learner self-repair

The importance of self-repair processes had already been the subject of investigation back in 1977 when Schegloff, Jefferson and Hacks pointed out that L1 conversation provides centrally for self-repair over other-repair. Kasper (1985) studied repair in foreign language teaching and concluded that correction patterns are determined by different learning and teaching goals and that self-initiated, self-completed repair is preferred by both learners and teachers. Green and Hecht (1993) looked at self-correction in English of both native and non-native (German) high school students and the results of their research throw a favorable light on the efficacy of this strategy.

More recently Lyster and Ranta (1997: 57) pointed out the importance of student-generated repairs for L2 learning for at least two reasons:

- a. they allow opportunities for learners to automatize the retrieval of target language knowledge that already exists in some form (e.g., as declarative knowledge, cf. Hulstijn 1990) and
- b. when repair is generated by the learners, they draw on their own resources and thus actively confront errors in ways that may lead to revisions of their hypotheses about the target language (Pica et al. 1989; Swain 1995).

We analyzed the self-repair behavior of this group of EFL learners to see if it could be a source of positive input. The oral production in both the L-L and the L-NS dyads was coded across tasks. The oral production of the seven L-L and the seven L-NS dyads in the four communication tasks was analyzed in order to identify the number of non-target language utterances (NTLU) against the backdrop of the total number of utterances. This counting was done in order to have a sense of proportionality of the repair processes (self- and other) under study. All utterances were coded as either having an error or not. Hesitations and false starts were excluded in the counting of utterances without errors. Table 1 features the classificatory matrix for NTLs and Table 2 how the errors were categorized:

Table 1. Classificatory matrix for NTLUs

Error utterances	Repaired	Self-repaired
		Other-repaired
	Unrepaired	

Table 2. Error categorization

Errors	Phonological	
	Lexical	
	Morphosyntactic	Subject
		Preposition
Other		

As Table 1 shows, if an error is made, then that error can be either repaired or not. In the latter case, the outcome will be that the error remains. However, if it is repaired, the learner can correct it himself/herself (self-repaired) or a peer can offer the repair (other-repaired). As for the classification of errors, we follow here the classificatory matrix in Lyster and Ranta (1997) but have expanded on the morphosyntactic category. Within “others” the following sections were considered (see Table 6): article usage, verbal form, word order, agreement and adverb placement

Table 3 shows the total number of utterances and of NTLUs in both L-L and L-NS interaction:

Table 3. Total number of utterances and NTLUs in L-L and L-NS interaction

	L-L		L-NS	
Number of utterances	Task 1:	466	Task 1:	527
	Task 2:	1592	Task 2:	975
Total		2058		1502
Total number of NTLU		170 (8.26%)		20 (1.33%)

Out of a total of 170 NTLUs in the L-L dyads, only 33 (19%) undergo some kind of modification by the speaker or his/her interlocutor but 137 (81%) go completely unrepaired. A similar picture was found in the L-NS dyads: out of a total of 20 NTLU, only 5 (25%) undergo some kind of modification by the speaker or his/her interlocutor but 15 (75%) go again completely unrepaired. The difference in the actual numbers corresponding to the L-L and the L-NS dyads is statistically non-significant ( $p$ -value = 0.55552).<sup>4</sup> The first observation that should be made is that the overwhelming majority of learner NTLUs goes unaddressed both in L-L and in L-NS interaction.

Another issue that should be mentioned in relation to the figures presented in Table 3 is that the number of utterances in the second task – decision making – is always higher than that in the first task – information gap – in both the L-L and the L-NS dyads. The vignettes in the picture tasks seemed to have helped the resolution of the stories, whereas the decision making tasks – specifically the one on survival materials –, with no accompanying visual aids, were more cognitively challenging for the students (Pica and Doughty 1986; Pica et al. 1989, 1991). There is also a significant difference ( $p$ -value = 0.000001) between the number of NTLUs between the L-L and the L-NS dyads (170 vs 20). A possible explanation was provided by post-task-interviews with the seven learners who participated in both L-L and L-NS dyads. They mentioned that the familiarity with their interlocutor in the L-L dyads made them feel more at ease and that is probably why more NTLUs were found (Gass and Varonis 1984). Learners tried to be more careful when interacting with the native speaker and avoided linguistic areas in which they were aware of their problems.<sup>5</sup>

Let us now consider the distribution of repaired utterances in both L-L and L-NS dyads:

Table 4. Percentages of repaired utterances in L-L and L-NS dyads

L-L dyads		L-NS dyads	
Repaired utterances:	19% (33/170)	Repaired utterances:	25% (5/20)
1. Self-repaired:	64% (21/33)	1. Self-repaired:	60% (3/5)
2. Other-repaired:	36% (12/33)	2. Other-repaired:	40% (2/5)

Out of the very few utterances that underwent some kind of repair on the part of the speaker or the interlocutor, we observe that this repair was mainly confined to self-repair in both L-L and L-NS dyads, with no significant difference between the two groups.<sup>6</sup> This finding is consistent with what



happens in L1 conversational interaction, that is, we seldom other-correct because we assume that our interlocutor will self-correct (Schegloff, Jefferson and Hacks 1977) and also with findings in L2 classrooms (Kasper 1985).

Table 5 features the types of repaired utterances in both L-L and L-NS dyads:

Table 5. Repaired utterances in L-L and L-NS interaction

	L-L dyads	L-NS dyads
REPAIRED UTTERANCES		
<b>Self-repaired:</b>	<b>64 % (21/33)</b>	<b>60 % (3/5)</b>
Phonological	14 % (3/21)	0
Lexical	10 % (2/21)	40 % (2/5)
Morphosyntactic	76 % (16/21)	20 % (1/5)
<b>Other-repaired:</b>	<b>36 % (12/33)</b>	<b>40 % (2/5)</b>
Phonological	17 % (2/12)	0
Lexical	58 % (7/12)	0
Morphosyntactic	25 % (3/12)	100 % (2/2)

In L-L interaction the higher percentage of learner self-correction (76%) goes to morphosyntactic errors and almost half of these (37%) are errors that have to do with third person singular possessive adjectives or direct object pronouns as in the following example:

(1) *It is at night and another man is coming and hits her ... him.*

Examples of phonological and lexical self-repair are given in (2) and (3) respectively:

(2) *Dry fruits ([fru:its]) ... fruits ([fruits]), for example, or fresh vegetables...*

(3) *But we are human beings, we are made of flesh and blood and just do... make mistakes.*

In L-NS interaction the higher percentage of errors that undergoes self-correction is featured by lexical errors (40%):

(4) [...] *I would definitely not take... give help to the second one because...*

In what follows examples are provided of other-repaired utterances in the different categories illustrated in Table 5:

Table 6.

<b>L-L dyads</b>		
Phonological		
(5)	Learner A	Learner B
	yeah, me too, probably calcium and sugar, what about sugar? flour ([flauə]), ok.	I ... no, no sugar, I would prefer flour ([flour])
Lexical		
(6)	Learner A	Learner B
	Ok, like alcohol, because it is necessary if you don't have ... eh ..., first auxiliaries ... yeah, and fresh water.	first aid ...
Morphosyntactic		
(7)	Learner A	Learner B
	So, what about a tent? A tent will be necessary ...  we agree, we agree (with emphasis) about that... then we take a tent... you said blankets?	yes, that's ... we are agree we are agree...
<b>L-NS dyads</b>		
Morphosyntactic		
(8)	Learner	Native speaker
	[...] but it could be significant the fact that he is not ... she is not constant	that she (with emphasis) is not ..

However, all the percentages in Table 5 have to be considered within the context of the proportions provided above. That is, although the percentage

given for learner self-correction of lexical errors in L-NS interaction is, as just mentioned, 40% the actual number of lexical errors self-corrected is just two. The infrequency with which learners self- and other-repaired made these strategies unlikely as a source of modified output. That is, learners provided a context for self- and other- repairs but, when those opportunities are analyzed in detail, we see that they are not really significant, even more so when we compare the self- and other-repaired figures with those of the utterances that go unrepaired. Table 7 provides this information:

Table 7. Percentages of unrepaired utterances in L-L and L-NS interaction

	L-L dyads	L-NS dyads
UNREPAIRED UTTERANCES		
<b>Total</b>	<b>81 % (137/170)</b>	<b>75 % (15/20)</b>
Phonological	12 % (16/137)	0
Lexical	11 % (15/137)	13 % (2/15)
Morphosyntactic	77 % (106/137)	87 % (13/15)
Subject	23 % (25/106)	8 % (1/13)
Preposition	21 % (22/106)	8 % (1/13)
Others: <sup>7</sup>	56 % (59/106)	84 % (11/13)

Most of the utterances that go unrepaired are largely confined to subject anaphora and the misuse of prepositions, together with article over- and under-suppliance and verbal tense and aspect features. This finding is also consistent with Bardovi-Harlig and Bofman (1989) and Storch (1998) in which this type of errors are mentioned as being persistent areas of concern for advanced learners. Examples of each category follow:

Phonological

- (9) *If we don't take salt ([ 'sa:lt]) the food would be very insipid or without taste ....*

Lexical

- (10) a. *[...] but being on duty for 27 hours has to be very, very tired ...*

Morphosyntactic

- (11) Lack of subject pronouns (indicated by in the text)  
(learners are trying to provide the right order in a series of vignettes)

Learner A

*ah! we can put this before so  
we can put 'h' before 'e' and  
'f' and ... he stands up  
again here?*

*and the last one I have is the  
letter 'c', \_\_\_\_ is this man, the  
main character, in pain, totally  
in pain...*

Learner B

*yes, and \_\_\_\_ is the first one*

- (12) Preposition

Learner A

*ok  
well, it depends of the extra clothes*

Learner B

*these two things then. The third  
one will be coats and jackets or  
extra-clothes*

Others

- (13) Verbal forms

Learner A

*and ... what else? I don't  
think ... eh ... you have said  
flashlight?*

Learner B

*yeah, because if you see a ship  
you can held one and say, hey,  
we are here!*

The modifications made by the learner and his/her interlocutor enable them to focus on form-meaning relationships but, clearly, considering the amount of NTLU that goes completely unrepaired we cannot conclude by saying that this strategy is helping learners to focus on those form-meaning relationships that are crucial to their morphosyntactic development.<sup>8</sup>

### 3.2. Completion strategies

Completion is a kind of scaffolding that has been identified in research on the collaborative dialogue that takes place between two learners (Pica and Doughty 1985; Pica et al. 1995; Swain 1995). Although it can be manifested in a variety of ways, completion is characterized by one interlocutor's hesitation over a word or sentence constituent, and the other interlocutor's suggesting the missing item.

Completions are a source of positive input (not modified input) because the completion makes the listener focus on the meaning of the speaker, then supply the form to map onto that meaning. By means of this type of scaffolding, learners in the present study were observed to offer appropriate words or phrases in order to complete each other's utterances.

We found two types of completion in our database:

- a. completion of interlocutor A's utterance: interlocutor B supplied word or phrase
- b. completion of interlocutor A's utterance + connector: interlocutor B supplied word or phrase

An example of each type follows:

- a. completion of interlocutor A's utterance:

(14) interlocutor B provides word:

Learner A	Learner B
<i>a tent, yes, because if it rains, and we cannot find any, you know, any hole or any ...</i>	<i>construction</i>

(15) interlocutor B provides phrase:

Learner A	Learner B
<i>if it gets wet, it doesn't ...</i>	<i>right, it doesn't ... it is not worth it</i>

b. Completion of interlocutor A's utterance + connector:

(16) interlocutor B provides word:

Learner A	Learner B
<i>coffee or tea, then, coffee or tea, whiskey and ...</i>	<i>fresh water</i>
<i>fresh water</i>	

(17) interlocutor B provides phrase:

Learner A	Learner B
<i>[...] and then ...</i>	<i>he stands up</i>

In examples (14) and (16), Learner B provides the lexical item his interlocutor needs to be able to continue their interaction. In (15) Learner B provides a phrase to complete Learner A's utterance and, once he has done that, he goes on with the conversation. In (17) Learner B provides a phrase which allows Learner B to continue with the description of the vignette in the corresponding task.

Our data revealed that completions constituted 9.2% (191) of the total number of utterances (2058) in the L-L dyads versus only 0.73% (11) of the total utterances of the L-NS dyads (1502). Thus, we found almost twelve times the number of completions among the learners when interacting with each other. The direction of this finding is consistent with Pica and Doughty (1985) and Pica et al. (1995) and shows that this strategy seems to be really helpful when two learners are interacting: the use of completion, as we have seen in the above examples, allows for further exchange of messages and moves the discourse along. However, completion of the message by their peers is mainly focused on the provision of lexical items and not much syntax building is present. This finding is similar to what Williams (1999: 609–10) reports in her descriptive study with eight learners from four different proficiency levels.<sup>9</sup> Once more, advanced learners do not seem to make a lot of effort to adjust their production above the word/phrasal level since their messages are quite comprehensible and allow their interaction to flow without major interruptions.

#### **4. Some pedagogical implications**

With respect to our first research question, we have seen that advanced learners do indeed engage in interactional modification of their discourse by means of strategies such as repair (mainly self-repair) and collaborative discourse. Our second research question considered the issue of whether or not these modifications address the learners' needs for positive and negative input and production of modified output. In principle, self-repair and other-repair provide a context for modified output but the infrequency with which learners used these strategies make them unlikely as a source of modified output. Completions, on the other hand, provide positive input but, in an important number of cases, the learners are just supplying a word or a phrase and not much syntax building is provided.

As for the importance of these strategies to enable learners to focus on L2 form-meaning relationships (research question 3), self-repair draws learners' attention to form in the sense that the learner recognizes that s/he has given an imprecise form to the meaning conveyed. However, considering the amount of NTLU that goes unrepaired, we cannot conclude by saying that this strategy is helping learners to focus on those form-meaning relationships that are crucial to their morphosyntactic development (research question 4). Completions make the listener focus on the meaning of the speaker's message so that s/he can supply a form to map onto that meaning but in the database analyzed completions mainly follow word searches at the end of a phrase.

Results from this study reinforce the conclusion reached by a number of researchers (Celce-Murcia 1991; Fotos 1998; Spada and Lightbown 1989; Williams 1999, 2001) that communicative language teaching alone is not sufficient to promote high levels of accuracy in learners.<sup>10</sup> This is one of the reasons for the current interest in reexamining the issue of form-focused instruction in the classroom where the primary focus is on communication (Doughty and Williams 1998; Ellis 2001). If we concentrate on what actually happens when different types of tasks are carried out, if we carefully analyze the quality of the language in those tasks, we will realize that different types of task set up different patterns of language use (Bygate 1999; Williams 1999). The language features identified can be later explored and exploited in order to target specific problematic areas.

By looking at the interlanguage of this advanced group of learners, we should think about possible tasks that could promote repair and completion strategies so that they are targeted on specific form-meaning relationships characteristic of this group, specific grammatical areas that are clearly overlooked when learners simply converse over communication tasks.

As Green and Hecht (1993: 161) suggest, learners can be encouraged to develop the strategy of self-correction, for example. From a pedagogical point of view, they point out, self-correction “may be seen as part of an education for autonomous learning”. But, how can this idea of encouraging self-correction be implemented in the foreign language classroom? One suggestion is to confront learners with oral/written texts with unidentified errors, which can be identified by the learners themselves in group, pair or individual work. Although, as Aston (1986) already noticed, we do not want learners to be constantly self-correcting because the flow of conversation will be seriously altered, “long-term language teaching cannot afford to undervalue a striving for linguistic correction because it can often be an important factor in social acceptance by native speakers” (Green and Hecht 1993: 161).

From the analysis of the data we have seen that one of the areas of grammar that seems challenging and problematic for students even at this advanced level is pronoun reference. What kind of task could target this troublesome area? Recent research by Bygate (1999) shows that, although a task as a whole may obviously give rise to unpredicted language features, there is considerable evidence that task structure affects learners’ selection of features of language. In order to target the area of pronoun reference we could use, for example, a form-focused interactive task such as a jig-saw strip story narrative in which each of the learners participating holds part of the information necessary in order to complete the task. The rationale for the use of a narrative comes from the observation that the structure of narratives lends itself to extended discourse in which, starting from a picture prompt, the participant needs to provide a structure, a situation and use linguistic expressions to identify elements such as characters, places and activities. S/he requires attention to reference features in the discourse. In fact, in Bygate’s research (1999: 204), the narrative task is the one that “stretches the speakers more in terms of complexity of syntactic and lexical processing” in contrast with an argumentation task.

A second type of task that could be used to make learners aware of the gap between their production and the target language is a form-focused high input task such as dictogloss (Wajnryb 1990). A dictogloss is a reconstruction activity whose basic procedure consists in the learner simply listening to or reading a short text once or twice in its entirety and reconstruct it from memory (individually or in pairs/groups). The reconstructed text is then compared with the original. The dictogloss is designed to draw the learners’ attention to language form. It is claimed that during the co-construction of the passage, the students come to notice their grammatical strengths and



weaknesses and try to overcome these weaknesses when attempting to co-produce the text (Kowal and Swain 1994; Lapierre 1994; Nassaji 2000; Swain 1998; Swain and Lapkin 2000, 2001; Thornbury 1997).<sup>11</sup>

## **5. Conclusion**

The present study was motivated by the interest in considering whether EFL learners' needs at advanced levels were addressed through interaction other than negotiation of meaning. Results of our research show that during their interaction advanced learners were resourceful in using strategies such as repair (self- and other-) and completion. They were able to draw from their interlanguage store both to correct and clarify message meaning and to complete each other's messages. Both strategies allowed for further exchange of messages, moved the interaction along and seemed to facilitate the development of grammatical and lexical features.

However, in spite of years of exposure to meaningful and comprehensible input and opportunities for interaction in communicative classrooms, the numerous morphosyntactic imprecisions observed in the learners' output strongly suggest the need to devise form-focused tasks that target the specific grammatical and lexical features that still need to be mastered at this stage (see García Mayo 2001a, 2001b, 2002).

## **Notes**

1. Financial support in the form of grants UPV #103.130-HA087/97 and UPV #103.130-HA011/99 from the University of the Basque Country is hereby gratefully acknowledged. I wish to thank Professor Teresa Pica (University of Pennsylvania) for her many helpful suggestions on earlier drafts of this paper and to two anonymous referees for pointing out interesting issues. Thanks also go to Professor Vicente Núñez Antón (Department of Econometrics and Statistics – University of the Basque Country) for his assistance with the statistical analysis of the data and to the students who volunteered for this study. All errors remain my responsibility.
2. ERASMUS is the name of an exchange program established between different European universities. The program allows students to complete part of their degrees in different host universities within the European Community.
3. It was difficult to find second-year students who had TOEFL scores in the required range and that was the reason why seven out of the fourteen volunteers

were the ones taking part in both the L-L and the L-NS dyads. Section 2. 2 provides more details about the different content in the tasks in which these seven students participated. The comparison between L-L and L-NS dyads was established because in an EFL setting learners have limited access to input and feedback from NS teachers and to interaction with NSs outside the classroom. In other words, learners become each others' model for language learning (García Mayo and Pica 2000a, 2000b; Gass 1990). It was necessary to see if the learners' interactional behavior was similar/different when their partner was a native speaker.

4. Using a two-sample binominal test.
5. This is contrary to Takahashi (1989). In her study – placed within the socio-psychological perspective of Accomodation Theory (Giles and Smith 1979) – she found that Japanese learners became more hesitant and briefer when addressing a listener with their same native language background than when addressing a non-Japanese listener. They also reported feeling more uncomfortable.
6. Self-repair:  $p$ -value = 0.8752; Other-repair:  $p$ -value = 0.8628. However, one should consider the actual numbers to place the results in an appropriate context.
7. The distribution of the unrepaired utterances within this category in L-L dyads is as follows:
 

Article misuse:	25% (15/59)
Verbal forms:	20% (12/59)
Word order:	15% (9/59)
Agreement:	10% (6/59)
Adverb placement:	8% (5/59)
Others:	20% (12/59)
8. Although the figures are very low, there is a significant difference in the number of NTLUs and other-repairs elicited by task 2 (decision making) in L-L and L-NS dyads but no significant differences in the number of self-repairs. No significant differences were found either between mixed-gender dyads and like-gender ones for NTLUs, self-repairs or other-repairs, unlike those found by Gass and Varonis (1986) for negotiation of meaning.
9. In her own words: “The proportion of the lexically oriented LREs [language-related episodes] is about 80% of all LREs for all proficiency levels, indicating a pervasive phenomenon independent of proficiency”. (Williams 1999: 610). A language-related episode is defined in Swain (1998: 70) as “any part of a dialogue in which students talk about the language they are producing, question their language use, or other- or self-correct”.
10. The activities central to communicative language teaching programs, devised with the goal of providing learners with opportunities for authentic communication, encourage, however, fluent and creative use of language resources (Allright 1984; Brumfit 1984; Nunan 1989; Prabhu 1987).
11. But see García Mayo (2001a, 2001b, 2002) for different results.

## References

- Allright, Richard L.  
1984 The importance of interaction in classroom language learning. *Applied Linguistics* 11: 17–39.
- Aston, G.  
1986 Troubleshooting in interaction with learners: The more the merrier? *Applied Linguistics* 7 (2): 128–143.
- Bardovi-Harlig, Kathleen and T. Bofman  
1989 Attainment of syntactic and morphological accuracy by advanced language learners. *Studies in Second Language Acquisition* 11: 17–34.
- Brumfit, Christopher J.  
1984 *Communicative Methodology in Language Teaching*. Cambridge: Cambridge University Press.
- Bygate, Martin  
1999 Quality of language and purpose of task: Patterns of learners' language on two oral communication tasks. *Language Teaching Research* 3 (3): 185–214.
- Celce-Murcia, Marianne  
1991 Grammar pedagogy in second and foreign language teaching. *TESOL Quarterly* 25: 459–480.
- Chaudron, Craig  
1988 *Second Language Classrooms. Research on Teaching and Learning*. Cambridge: Cambridge University Press.
- Crookes, Graham  
1990 The utterance and other basic units for second language discourse analysis. *Applied Linguistics* 11 (2): 183–199.
- Doughty, Catherine and Jessica Williams (eds.)  
1998 *Focus on Form in Classroom Second Language Acquisition*. Cambridge: Cambridge University Press.
- Duff, Patricia  
1986 Another look at interlanguage talk: taking task to task In *Talking to Learn: Conversation in Second Language Acquisition*, Richard Day (ed.), Rowley, MA: Newbury House.
- Ellis, Rod  
2001 Investigating form-focused instruction. *Language Learning* 51: 1–46.
- Fotos, Sandra  
1998 Shifting the focus from forms to form in the EFL classroom. *ELT Journal* 52 (4): 301–307.
- García Mayo, María del Pilar  
2001a Are collaborative focus-on-form (FonF) tasks worthwhile among advanced EFL learners? In *Recent Perspectives on Discourse*, Ana. I. Moreno and Vera Colwell (eds.) León: Universidad de León (CD-ROM format).

- 2001b Focus-on-form tasks in EFL grammar pedagogy. In *Language Awareness in the Foreign Language Classroom*, David Lasagabaster and Juan Sierra (eds.), 221–236. Bilbao: Universidad del País Vasco.
- 2002 The effectiveness of two form-focused tasks. *International Journal of Applied Linguistics* 12 (2): 156–175.
- García Mayo, María del Pilar and Teresa Pica
- 2000a L2 learner interaction in a foreign language setting: Are learning needs addressed? *International Review of Applied Linguistics* 38 (1): 35–58.
- 2000b Interaction among proficient learners: Are input, feedback and output needs addressed in a foreign language context? *Studia Linguistica* 54 (2): 272–279.
- Gass, Susan
- 1990 Second and foreign language learning: Same, different or none of the above? In *Second Language Acquisition – Foreign Language Learning*, Bill Van Patten and James F. Lee (eds.), 34–44. Clevedon: Multilingual Matters.
- Gass, Susan M. and Larry Selinker
- 1994 *Second Language Acquisition: An Introductory Course*. Hillsdale, N.J.: Lawrence Erlbaum.
- Gass, Susan M. and Evangeline M. Varonis
- 1984 The effect of familiarity on the comprehensibility of nonnative speech. *Language Learning* 34 (1): 65–89.
- 1986 Sex differences in NNS/NNS interactions. In *Talking to Learn*, Richard Day (ed.), 327–351. Rowley, MA: Newbury House.
- 1994 Input, interaction, and second language production. *Studies in Second Language Acquisition* 16: 283–302.
- Gass, Susan, Alison Mackey and Teresa Pica
- 1998 The role of input and interaction in second language acquisition. *The Modern Language Journal* 82(iii): 299–307.
- Giles, H. and P. Smith
- 1979 Accommodation theory: Optimal levels of convergence. In *Language and Social Psychology*, H. Giles and R. St. Clair (eds.), London: Basil Blackwell.
- Green, Peter S. and Karlheinz Hecht
- 1993 Pupil self-correction in oral communication in English as a Foreign Language. *System* 21 (2): 151–163.
- Holliday, Lloyd
- 1995 NS syntactic modifications in NS-NNS negotiations as input data for second language acquisition of syntax. Ph.D. diss., University of Pennsylvania.
- Hulstijn, Jan
- 1990 Appealing to consciousness in the L2 classroom. *AILA Review* 11: 57–68.

- Kasper, Gabriele  
1985 Repair in foreign language learning. *Studies in Second Language Acquisition* 7: 200–215.
- Kowal, Maria and Merrill Swain  
1994 Using collaborative language production tasks to promote students' language awareness. *Language Awareness* 3 (2): 72–93.
- Lapierre, Donna  
1994 Language output in a cooperative learning setting: Determining its effects on second language learning. M.A. diss., University of Toronto, Toronto: Canada.
- Linnell, Julian  
1995 Negotiation as a context for learning syntax in a second language. Ph.D. diss. University of Pennsylvania.
- Long, Michael  
1996 The role of the linguistic environment in second language acquisition. In *Handbook of Second Language Acquisition*, William C. Ritchie and Tej K. Bhatia (eds.), 413–468. New York: Academic Press.
- Lyster, Roy and Leila Ranta  
1997 Corrective feedback and learner uptake. *Studies in Second Language Acquisition* 19: 37–66.
- Mitchell, Rosamond and Florence Myles  
1998 *Second Language Learning Theories*. London: Edward Arnold.
- Nassaji, Hossein  
2000 Towards integrating form-focused instruction and communicative interaction in the second language classroom: Some pedagogical possibilities. *The Modern Language Journal* 84 (iii): 241–250.
- Nunan, David  
1989 *Designing Tasks for the Communicative Classroom*. Cambridge: Cambridge University Press.
- Pica, Teresa  
1994 Research on negotiation: What does it reveal about second language learning conditions, processes, and outcomes? *Language Learning* 44(3): 493–527.
- Pica, Teresa and Catherine Doughty  
1985 Input and interaction in the communicative classroom: A comparison of teacher-fronted and group activities. In *Input in Second Language Acquisition*, Susan Gass and Carolyn Madden (eds), 115–132. Rowley, MA: Newbury House.  
1986 “Information gap” tasks: Do they facilitate second language acquisition? *TESOL Quarterly* 20 (2): 305–325.
- Pica, Teresa, Richard Young and Catherine Doughty  
1987 The impact of interaction on comprehension. *TESOL Quarterly* 21: 737–758.

- Pica, Teresa, Lloyd Holliday, Nora Lewis and Lynelle Morgenthaler  
1989 Comprehensible output as an outcome of linguistic demands on the learner. *Studies in Second Language Acquisition* 11: 63–90.
- Pica, Teresa, Lloyd Holliday, Nora Lewis, Dom Berducci and Jeanne Newman  
1991 Language learning through interaction: What role does gender play? *Studies in Second Language Acquisition* 11: 63–90.
- Pica, Teresa, Felicia Lincoln-Porter, Diana Paninos and Julian Linnell  
1995 What can second language learners learn from each other? Only their researcher knows for sure. *Working Papers in Educational Linguistics* 11 (1): 1–36.
- Prabhu, N. S.  
1987 *Second Language Pedagogy*. Oxford: Oxford University Press.
- Sadow, S. A.  
1982 *Idea Bank: Creative Activities for the Language Class*, Rowley, MA: Newbury House.
- Savignon, Sandra  
1991 Communicative language teaching: State of the art. *TESOL Quarterly* 25(2): 261–277.
- Schegloff, Emanuel, Gail Jefferson and Harvey Sacks  
1977 The preference for self-correction in the organization of repair in conversation. *Language* 53: 361–382.
- Spada, Nina and Patsy Lightbown  
1989 Intensive ESL programs in Quebec primary schools. *TESL Canada Journal* 7: 11–32.
- Storch, Neomy  
1998 A classroom-based study: Insights from a collaborative text reconstruction task. *ELT Journal* 52/4: 291–300.
- Swain, Merrill  
1985 Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In *Input in Second Language Acquisition*, Susan Gass and Carolyn Madden (eds.), 235–253. Rowley, MA: Newbury House.
- 1995 Three functions of output in second language learning. In *Principles and Practice in Applied Linguistics: Studies in Honour of H.G. Widdowson*, Guy Cook and B. Seidlhofer (eds.), 125–144. Oxford: Oxford University Press.
- 1998 Focus on form through conscious reflection. In *Focus on Form in Classroom Second Language Acquisition*, Catherine Doughty and Jessica Williams (eds.), 64–81. Cambridge: Cambridge University Press.
- Swain, Merrill and Sharon Lapkin  
2000 Task-based second language learning: The uses of the first language. *Language Teaching Research* 4 (3): 251–274.

- 2001 Focus on form through collaborative dialogue: Exploring task effects. In *Researching Pedagogic Tasks. Second Language Learning, Teaching and Testing*, Martin Bygate, Peter Skehan and Merrill Swain (eds.), 99–117. London: Longman.
- Takahashi, T.  
1989 The influence of the listener on L2 speech. *Variation in Second Language Acquisition Volume I: Sociolinguistic Issues*, In Susan M. Gass, Carolyn Madden, D. Preston and Larry Selinker (eds.), Clevedon: Multilingual Matters.
- Thornbury, Scott  
1997 Reformulation and reconstruction tasks that promote noticing. *ELT Journal* 51(4): 326–335.
- Ur, Penny  
1996 *Discussions that Work*. Cambridge: Cambridge University Press.
- Wajnryb, Ruth  
1990 *Grammar Dictation*. Oxford: Oxford University Press.
- Williams, Jessica  
1999 Learner-generated attention to form. *Language Learning* 49 (4): 583–625.  
2001 The effectiveness of spontaneous attention to form. *System* 29: 325–340.





# **Assessment of the role of communication tasks in the development of second language oral production skills**

*Peter Griggs*

## **1. Introduction**

One of the central problems facing instructed second language learning concerns the gap between the learner's internal rules, which he uses to guide his performance in the target language, and the external rules, constituting the norm, which language teaching attempts to impose on his performance. In order to carry out L2 classroom activities, pupils use both intuitive rules acquired implicitly and incidentally from their experiences in the language, and rules of differing degrees of explicitness constructed on the basis of the pedagogical grammar to which they are exposed. In an orthodox communicative approach, language activities which are initially controlled and formal are destined in later stages to become freer and more functional, and increasingly closer to real-life communicative activity. The learning problem can therefore be considered to be twofold: firstly, how language input which is structured and presented in terms of the teacher's pedagogical grammar is used by the learner to construct an internal grammar; and, secondly, how the latter is progressively adapted to increasingly communicative and authentic language needs.

The aim of this study is to assess the role that communication tasks play in developing oral skills in such a learning context. The notion of communication task used in the study corresponds to Nunan's definition (1989: 10): "a piece of classroom work which involves learners in comprehending, manipulating, producing or interacting in the target language while their attention is principally focused on meaning rather than form". Ideally, communication tasks place learners in the centre of the learning process by creating an interactional framework in which they solve language problems in order to fulfill communicative needs. Oral task work is generally characterised by the organisation of classroom activity in groups or in pairs and by the creation of a gap between different sources or forms of information, which serves as a trigger for the learners to partake in communicative

interaction<sup>1</sup>. Two main arguments can be put forward in favour of the use of communication tasks in the context of second language teaching. Firstly, they provide a favourable context for learning: the organisation of task work in small groups permits a greater amount of language production; the learner is more autonomous and more involved in the learning process; the communicative needs created by tasks are likely to increase a learner's level of motivation; group work can help generate a positive affective climate in the class. A second argument concerns their role in bridging the gap between language learning in the classroom and second language use in real communicative situations. Nunan (1989) advocates not only the use of authentic tasks based on real life situations (role plays, simulations) but also that of pedagogical tasks which allow learners to develop the communicative skills required to function in the real world.

The corpus used in the present study is made up of recordings of various types of oral communication task performed by intermediate learners of English – a group of young adults on a professional training scheme and a group of non-specialist first-year students – all of whom are French native speakers<sup>2</sup>.

## 2. Theoretical considerations on task-based learning

### 2.1. Limits to the interaction hypothesis

In the literature (eg. Duff 1986; Long & Porter 1985; Pica 1987, 1994; Varonis & Gass 1985) both native/non-native speaker interaction and second language communication tasks between learners have often been studied in the perspective of what has been called the *interaction hypothesis* (Ellis 1990: 107–117; Long 1983; Pica 1994), according to which the negotiation of meaning sets off natural processes of acquisition. Pica (1994: 497) defines negotiation as: “a process in which a listener requests message clarification and confirmation and a speaker follows up these requests, often through repeating, elaborating, or simplifying the original message”.

Viewed in the framework of Krashen's *input hypothesis* (Krashen 1982), negotiation and the modified speech it entails (grammatically less complex utterances, higher-frequency vocabulary, avoidance of idiomatic expressions) were initially considered to be an efficient means of providing the non-native speaker with comprehensible input (Long & Porter 1985). More recently, the benefit of the negotiation of meaning has been assessed more in terms of the attention to form it induces as learners attempt to process

meaningful input, reflecting the increasingly consensual view among second language researchers that learners need to notice structural features of the language in order to assimilate them into their interlanguages (Schmidt 1990, 2001; Sharwood Smith 1993; Tomlin & Villa 1994). It is also argued that the communicative goals that non-native speakers pursue during verbal interaction may push them to modify their own output in order to enhance their performance and thereby engage in a higher level of linguistic processing and try out new hypotheses concerning the target language (Pica 1994; Pica et al. 1989; Swain 1985).

Now while this communicative perspective may be appropriate for interactions between native and non-native speakers or between non-native speakers with different first languages or of different levels of second language ability, it appears to be less so for interactions involving intermediate learners sharing the same native language. Indeed an analysis of the corpus used in the present study has revealed that communication problems requiring the negotiation of meaning are somewhat limited and that consequently learners receive a relatively small amount of new input data (Griggs 2000). This can be explained firstly by the fact that the gap between the learners' respective language systems is small and that the learners can always resort directly or indirectly to their common first language. Furthermore, the information gap underlying the task does not necessarily lead learners to focus on form in order to negotiate meaning. In this connection, Skehan (1996, 1998) evokes the risk, from an acquisitional point of view, of learners resorting too systematically to a lexical mode of communication, based on prefabricated patterns rather than on analysed structures, and of proceduralising prematurely the solutions they have found to overcome communication problems. The danger is compounded by the absence of an expert speaker acting as an authority to ensure that the norms of the target language are respected. As for the reciprocal scaffolding carried out by the two learners during the interaction, it runs the risk of getting stuck in a communicative mode, where the partners content themselves with confirming mutual understanding, rather than working collaboratively to improve their mastery of the target language (Nussbaum et al. 1999). These problems can be illustrated in these two examples from the corpus <sup>3</sup>:

- (1) J it's a woman eh: standing eh in front of the buffet.  
B yes,  
J and: O me they she wear, eh: a hat. no.  
B yes. it's: O she has O eh this one then. eh see on the right, or on the left. on your picture.

- J on the left.  
 B on the left. me too.  
 J he has  
 B oh, how many button on his coat,  
 J one two three four five,  
 B one two three four me four, O yes there's a problem

In the first example, the task requires the two learners to find the differences between two almost identical pictures by describing their own picture without looking at their partner's. They accomplish this part of the task in an efficient manner, establishing three similarities (the presence and the location of a woman in the two pictures and the fact she is wearing a hat) and one difference (the number of buttons on her coat). However, the descriptions the learners produce in carrying out the task are far from elaborate; the utterances are short and elliptic ("me too", "see on the right", "he has"), characterized by poor and sometimes incorrect syntax ("it's a woman", "how many button on his coat", "she wear") and relying heavily on a few key lexical phrases ("on the left", "on the right", "there's a problem"). Consequently the task in no way pushes the learners to work on the quality of their language production.

- (2) G my favourite 0 **matière, je ne sais pas le dire**  
 J **matière**  
 G ok my favourite **matiere** is sport

In the second example, G resorts to the shared native language in order to solve a lexical problem, while admitting explicitly that he doesn't know the equivalent word in English ("**matière, je ne sais pas le dire**": 'matière I don't know how to say it'). This could be interpreted by his partner as an appeal for help to search for the missing word. By repeating the borrowed French word, however, the latter indicates both that he has understood the utterance and that he doesn't intend to supply the missing word, giving thus his approval to a form of communication that accords priority to mutual understanding over correct target language use. What counts is communication at all costs.

## 2.2. A cognitive perspective of second language learning

The communicative perspective tends to presuppose paradigms which are based on dichotomies between incidental, unconscious acquisition and intentional, conscious learning (Krashen 1982; Zobl 1995) and/or between the acquisition and analysis of knowledge and the development of control (Bialystok 1990; Bialystok and Sharwood Smith 1985; Ellis 1994). These paradigms attribute primary importance to implicit processes of acquisition, triggered by positive input, and thus relegate explicit or controlled processes, and learner output, to a peripheral role. Formal instruction and explicit learning are generally considered to have only indirect and often delayed effects on interlanguage development. For example, according to the *weak interface position* presented by Ellis (1994: 88–89) “explicit knowledge derived from formal instruction may convert into implicit knowledge, but only if the learner has reached a level of development that enables her to accommodate the new linguistic material” and formal instruction may also be “a way of helping learners develop greater control over L2 knowledge (explicit or implicit) that they already possess”. Similarly, Schmidt (2001: 10) claims that “one major role of explicit instruction is that, by changing expectations, it helps focus attention on forms and meanings in the input, a prerequisite for subsequent reprocessing”. As for the role of output in second language acquisition, even the proponents of the *comprehensible output hypothesis* (Swain 1985; Swain and Lapkin 1995) regard it as being only of secondary importance compared to that of input, and seem to define its effect on interlanguage development in terms of the dual-processing model previously mentioned in which the initial output is based on implicit knowledge (Swain and Lapkin 1995: 375): “It might be that producing language forces learners to recognize what they know or know only partially. This may trigger an analysis of incoming data, that is, a syntactic analysis of input, or it may trigger an analysis of existing internal linguistic resources, in order to fill the knowledge gap”.

The position adopted in the present study seems to me to be different insofar as it doesn't assume separate systems of knowledge. The claim that is made is that metalinguistic activity carried out during communicative second language production has a direct effect on the construction of second language skills. Metalinguistic activity is defined in this case as increased attention to the manipulation of language form during communicative interaction and is considered to be overtly displayed in self and other repair and in certain types of recourse to the first language.

The claim is supported by the results of a previous longitudinal study (Griggs 1997) in which six pairs of French intermediate non-specialist students of English performed six communication tasks at two-week intervals over a university semester. It was shown that a group of learners characterised by a high rate of metalinguistic activity during task performance made significantly more progress over the semester in terms of accuracy, and slightly more in terms of fluency, than a group whose rate of metalinguistic activity was low. The tasks were of two types, alternating between three discussion tasks based on a questionnaire or a list of controversial statements and three tasks involving interview simulations. After preparing individually a task for ten minutes, the pairs of students recorded their verbal interactions in a language laboratory with their microphones plugged into the same recorder. They were not allowed to stop their recorders during the ten-minute recording period. The students were then asked to listen to their recordings together and to note down and correct the mistakes that they noticed. It was thought that the language laboratory setting and the post-recording metalinguistic assignment would reinforce the students' perceptions of being in a learning framework.

By calculating the rate of repair work carried out by each learner in relation to the number of words he or she produced, it was possible to separate the learners into two equal groups according to the frequency of their metalinguistic activity: group 1 had an average rate of 1 repair per 33 words and group 2 an average rate of 1 repair per 70 words, which a t-test showed to represent a significant difference between the two groups ( $p < 0.01$ ). Progress in the two types of task was assessed separately by comparing the first and last task of each type. Fluency was measured in terms of the number of words produced per minute and accuracy was based on the rate of error per number of words. In both types of task, group 1 showed slightly higher progress in fluency than group 2: the average rates of progress were respectively 1.34 and 1.25 compared to 1.11 and 1.06, which was not however significant according to a t-test. As for improvement in accuracy, group 1 significantly outperformed group 2 in both task types, with average rates of progress at 1.42 compared to 0.95 ( $p < 0.05$ ) and at 2.00 compared to 0.86 ( $p < 0.025$ ).

The results of this study would seem to imply that the contribution of communicative tasks to language learning depends not so much on the framework they offer in activating natural processes of acquisition as on the cognitive strategies learners themselves adopt while performing them. Such learner strategies should be distinguished from communication strategies as they have been defined and described in the literature devoted to

them (eg. Bialystok 1990; Faerch and Kasper 1983) and which are generally considered to be a means for learners to overcome problems of communication, either by modifying or abandoning communicative aims or by devising alternative verbal or non-verbal plans to achieve them. On the contrary, successful learners will be the ones who attempt to use their language resources optimally while pursuing their communicative aims, who work collaboratively to solve language problems as they arise, and who may consciously exploit the communicative activity they are involved in to improve their mastery of language structures. The learning that is potentially induced is essentially output-driven rather than input-driven, involving the development of both fluency and accuracy skills through the collaborative construction of second language discourse. These characteristics are displayed in the following example:

- (3) E and do you wish you had more time to think before speaking.  
S OO **le temps de penser**, before speaking, 0 **attends** 00 yes I wish eh I had more time to think before speaking, because eh O I have eh time to to eh  
E prepare your answer.  
S prepare your answer. and to understand the answer and to, eh to eh **trouver mince**  
E find  
S **quoi**  
E to find OO **trouver** to find.  
S **oui oui** to find the good way, eh solution, of of eh this question.

In this example, the metalinguistic activity is initially triggered by the presence of a complex grammatical form “do you wish you had” in the questionnaire accompanying the task, but it then goes on to focus more on output than on input processes. Admittedly, S takes a certain time to understand and to reply to the question that her partner reads out to her. Nevertheless, her translation into her first language of part of the utterance (“**le temps de penser**”) seems to constitute as much a springboard to reproduce the model as a way of facilitating comprehension. Her partner then helps her to construct her reply by completing her utterance (“I have time to to eh” → “prepare your answer”) and by supplying her with a lexical item (“find”) in reply to an appeal formulated in the first language (“**trouver mince**”: ‘**trouver** damn it’). After repeating the word she was looking for, S finally reformulates part of her own utterance (“find the good way eh solution of of eh this question”). This example demonstrates the importance in commu-



nicative task work of learner collaboration focussing on interlanguage production and on the tuning of this production to target language norms. Starting from a standpoint of reciprocity based on their shared first language, the learners treat the second language not as an obstacle that separates them but as an object to construct together.

This corresponds to a cognitive view of learning, following a constructivist tradition, where learning takes place through problem-solving in the framework of goal-oriented activity, and in which implicit and explicit knowledge, controlled and automatic processes interlink synergetically in a single complex system based on general learning mechanisms (eg. Anderson 1982, 1983; Karmiloff-Smith 1992; McLaughlin 1987; Schneider and Schiffrin 1977).

The empirical study that follows is set in the framework of a model of second language acquisition which draws its inspiration primarily from Anderson's cognitive theory of learning (1982, 1983). The model, presented in greater detail in Bange, Carol and Griggs (2002), is based on the classic distinction in cognitive psychology between declarative knowledge, composed of facts required for generating actions, and procedural knowledge, composed of procedures required for performing them. The acquisition of a second language skill is considered to entail an evolution from declarative mental representations to procedural ones. Declarative knowledge is a reservoir of all the facts we have recorded in memory. In the context of second language learning, it includes not only grammar rules but also knowledge of different degrees of explicitness accumulated through direct experience with the second language or resulting from mastery of a first language. In order to generate performance, declarative information is assembled and converted into what Anderson calls *production rules*, which are practical inference rules consisting of a condition followed by the action that the condition entails: if X then Y. A production rule may be formal in nature: eg. "If you want to mark formally the past tense of a verb, then use the base form and add the suffix -ed"; or it may be a functional rule, based more on lexis than on syntax, of the type "If you want to thank somebody, then say 'thanks'".

The model envisages two phases of learning.

At an *initial stage*, the learner constructs a production rule from units of declarative knowledge using general problem-solving procedures, such as inference and analogy, in order to carry out a language act corresponding to the aim he has set himself. The initial rule or form is often approximate as regards the target language norm owing to the poverty of his interlanguage or the cognitive constraints imposed by the situation. The construction of



production rules will often be based, in the initial stages of second language acquisition, on analogies with first language knowledge, composed not only of the system of rules for communication on which the learner can model his second language production, but also on the semantic system of reference by which he comprehends the world<sup>4</sup>.

In the ensuing *evolution phase*, two parallel and often contradictory processes are in play. Firstly, a process of proceduralisation takes place as different production rules are synthesized into a single rule and their activation comes to rely less and less on declarative information in working memory. At the same time, interlanguage production is progressively tuned to target language norms. Tuning involves three types of learning process: firstly, *generalisation*, by which production rules become broader in their range of application, as the learner seeks similarities between rules and creates new rules that incorporate the features they have in common. In the course of verbal interaction, generalisations are likely to be promoted by positive feedback. The second process is *discrimination*, whereby a learner limits the range of a production rule, after comparing variables in correct and incorrect applications. Discrimination results either from negative feedback which is in contradiction with a production rule used by a learner or from his realising, through monitoring, that an utterance he produces based on memorised input data or on internal computations conflicts with a procedural rule he uses habitually. The third type of process is *strengthening*, by which rules which have been attested are consolidated through further use and rules which have not are weakened and eventually discarded.

### 3. An empirical study

In the empirical study that follows the same corpus was used as in the longitudinal study that has just been presented (Griggs 1997). The analysis focuses first of all on the second language performance of the most successful student in the previous study, Sandrine<sup>5</sup>, and the development of her use of two past tenses, the *simple past* and the *present perfect*, over the six tasks. In standard English the following distinction is generally made between the two tenses: the *simple past* refers to a definite time in the past which may be identified by a past time adverbial, the preceding language context or the context outside language; the *present perfect* is used to refer to states leading up to the present time, to indefinite events or habits in a period leading up to present time, and to past events with results in the present time. A speaker of French is accustomed to making a different distinction

in his native language between the *passé composé*, used for precise past actions which are limited in time, and the *imperfect*, used for past actions of undefined duration.

The choices of subject and language structure were motivated by two factors: firstly, Sandrine's second language discourse, and more particularly her use of the past tenses, is characterised by a large amount of self and other-repair, frequent recourse to first language mediation (metalinguistic comments and translation into and from the second language) and various types of hesitation phenomena (filled and unfilled pauses, repetitions, false starts, lengthening of syllables). Not only do these features reveal heightened metalinguistic reflection, they also provide the observer with clues as to the rules on which the learner is basing her productions. Secondly, a quantitative analysis of Sandrine's use of verbs in all contexts requiring either the simple past or the present perfect reveals a clear improvement in her mastery of these structures in the final task (see table 1).

Table 1. Quantitative analysis of Sandrine's use of past tenses

Tasks	N° contexts	Correct use	Functional error	Formal error
T1	1	1	0	0
T2	6	2	4	0
T3	0	0	0	0
T4	9	2	2	5
T5	1	1	0	0
T6	15	12	2	1

In this table, formal errors are those which do not conform to any morphological structure attested in the target language; whereas functional errors are those in which the structure used respects target language morphological rules but does not correspond to the tense that is required in the context. Tasks 1,3 and 5 are discussion tasks which, unlike simulation tasks 2, 4 and 6, give rise to practically no instances of past tense use.

### 3.1. Qualitative analysis of past tense use

The following analysis attempts to relate the improvement of Sandrine's mastery of the past tenses to the language work she carries out during task

performance and to explain this improvement in terms of the production rules she seems to be using<sup>6</sup>. It starts from the principle that the problematic tense is the present perfect, since, from the point of view of target language norms, this tense is overused by Sandrine and encroaches on the functional domain normally reserved for the simple past. An evolution in the functional distinctions Sandrine makes between the two tenses can be demonstrated by comparing examples from task 2 to examples from task 6.

3.1.1. Examples from task 2

- (4) S I have a lot of diplomas, O as  
E like  
S mm eh **alors** yes eh licence O licence of sociology, eh *I studied* psychology, psychology and sociology eh during eh my first year year of O eh *I have eh O I have had* a a degree, eh eh in eh eh O two years ago.
- (5) E and what are you diff different skill,  
S eh O *I worked eh O I have worked no I worked* before.  
E not not your experience, your skills
- (6) E what are your situation, situation family situation,  
E + or +  
S + yes + eh: I don't have a family, eh **oui** eh *I have lost* my parents eh O when I **quand j'étais** *I was* eh O enfant O *I was* children.
- (7) E *have you ever worked* in this sort of environment, OO  
S *I have* eh **j'étais** OO  
E is that your first job in this sort of environment with delinquency, or  
S *I live* al: eh *I live already* in this eh in this eh council eh I I know eh eh the child the eh this eh problem of the child, childrens eh

In these four sequences, the metalinguistic reflection carried out by Sandrine is abundantly manifest in the various hesitation phenomena (pauses, false starts etc.) which pervade her discourse, in the repetitions and self-repairs focussing on problematic structures (examples 4, 5, 6 and 7) and in the cases of translation to and from the first language (examples 6 and 7).

Sandrine's use of verbs in these sequences also shows that she is clearly making a distinction between the simple past and the present perfect. In

examples 4 and 6 the two tenses are used in the same utterances: “*I studied* psychology, psychology and sociology eh during my first year of O eh I have eh *I have had* a a degree, eh in eh O two years ago”; “*I have lost* eh my parents eh O when I **quand j’étais** I was eh O **enfant** O I was children”. In example 5, Sandrine hesitates between “*I worked* before” and “*I have worked* before” and finishes by opting for the former.

However, the choices being made do not always conform to target language norms. They appear to be based rather on a distinction between, on the one hand, short actions, limited in time, focusing on the result, expressed by the *passé composé* in French, and, on the other hand, actions or states of undefined duration, which focus on the activity and are expressed usually by the imperfect in French.

For the first function, Sandrine uses the present perfect in English: “*I have had* a degree two years ago” seems to be modelled on *j’ai eu une licence il y a deux ans*, the French verb *avoir* denoting a change in state which is not part of the semantic properties of the equivalent English verb ‘have’, but corresponds rather to the verb ‘get’; similarly, “*I have lost* my parents” can be considered to be modelled on the French *j’ai perdu mes parents*.

In the case of the three verbs expressing undefined duration and focusing on an activity or a state (‘study’, ‘work’ and ‘be’), Sandrine uses the simple past. “*I studied* psychology and sociology during my first year” would however require the *passé composé* in French, despite the intrinsic properties of the verb, as *j’ai étudié* would be limited in past time by the prepositional phrase *pendant ma première année*; “*I worked* before”, which according to target language norms, is less appropriate than the present perfect “*I have worked* before”, can be rendered either in the *passé composé* or in the imperfect in French; “when *I was* children”, which Sandrine explicitly translates from “**quand j’étais enfant**”, requires the use of the imperfect.

In example 7, Sandrine starts to reproduce the present perfect in reply to her partner’s question “*have you ever worked*”. She then hesitates, probably on realising that in French she would use the imperfect *j’étais*, which, according to the previous distinction, corresponds to the simple past. If one follows this reasoning, then her subsequent recourse to the simple present may constitute a strategy allowing her to avoid making the choice between the two past tenses.

In task 2, Sandrine seems therefore to be following past tense rules based on functional distinctions analogous to those operating in her first language and which could be formulated in the following manner:

*Rule 1:* If in French you use the *passé composé* then use the present perfect in English.

*Rule 2:* If in French you use the imperfect then use the simple past in English.

Although the occurrences of the two rules in task 2 are somewhat limited, their consistent and concurrent use implies that they have undergone a certain degree of generalisation and discrimination. Furthermore, in so far as Sandrine receives no negative feedback from her partner, the rules are likely to be strengthened and generalised to other contexts of past tense use. Indeed, in this sort of task-learning situation, unlike that of natural second language acquisition, it is improbable that Sandrine will receive sufficient positive or negative feedback to allow her to rectify defective rules and adjust them to target language norms. But nor does Sandrine put into use the past tense grammar rules she has no doubt been exposed to during her seven years' experience of English language learning. Instead, she constructs simpler, more accessible rules in order to be able to concentrate on the more important task of achieving immediate communicative goals.

### 3.1.2. Examples from task 6

- (8) E what kind of children, small children, or teenagers.  
S eh eh they are **oui** *they are* **non** *they have been*,  
E *they was* eh, how many children *there were* children *have they*  
S two children
- (9) S *I visited* eh  
E *I have visited*  
S **ah oui non non j'ai visité. il n'y a pas de temps. mais non** *I I visited* eh Canterbury in part of Eng of England which name is K name is Kent, eh *I see eh I saw* Westminster the the gigantesque cathedral, eh **voilà** eh and *I I have gone* eh **ça va arriver** *I went* eh in London and *I visited* some monuments and eh some museums eh
- (10) S yes yes, and *what have you been* eh **oui** *what have you been, why have you been* in USA  
E well it was a O an opportunity, a good opportunity, because the the the travel were cheap, and eh *I were* *I would like* *I want to go* here to see what is America

In task 6, confusion in the speakers' representations concerning past and present time is compounded by the nature of the simulation activity, in which learners imagine they have just spent a year in England or the USA. In example 8, uncertainty as to whether this experience is part of the past or part of the the present is reflected in the use of five verb forms involving three different tenses in a series of reformulations: "*they are*", "*they have been*", "*they was*", "*there were*", "*have they*". As Sandrine's reply, extending over two utterances, "*they have been* two children", doesn't conform to the previously established first language rule, one can postulate that she is attempting to apply the target language present perfect rule.

This movement away from first language modelling towards second language norms continues in example 9 which marks a turning point in the development of Sandrine's past tense rules. The three verb forms she uses ("*I visited*," "*I saw*", "*I went*") are in conflict with the first language based rule that has tended to underlie her production up to this point (see table 2), and conform rather to the target language rule whereby actions clearly situated in a past context require the simple past tense. This change is quite clearly accompanied by metalinguistic reflection. Sandrine's initial use of the simple past is wrongly corrected by her partner who proposes the present perfect form in its place, thus indicating his adherence to the same rule that Sandrine has been using until now. However, Sandrine repeats her initial utterance, and thereby insists on its correctness, after having translated it and attempted to justify her choice in French ("**ah oui non non j'ai visité. Il n'y a pas de temps.mais non**"). The meaning of the metalinguistic comment (literally 'there is no time') is unclear, but it may express her awareness that in English the way an action is situated in time is more of a determining factor for the choice of tense than the type and duration of the action. The two subsequent instances of simple past use follow on from self-repair, the second being accompanied also by another metalinguistic comment "**ça va arriver**" ('it's coming') expressing the cognitive effort being made.

In example 10, the last case of present perfect use "*why have you been in USA*", which according to the former rule would have been "*why were you in the USA*", is not really appropriate in the context as the use of the past participle 'been' indicates the interlocutor has come back from the USA, and therefore clearly situates this experience in the past. The misuse of the tense may result from the confusion over time caused by the task. It also indicates that Sandrine is attempting to employ a new target language rule limiting the use of the present perfect to actions related to the present and discriminating it thus from the simple past.

Task 6 is characterized therefore by a radical change and extension in the use of the simple past and a subsequent restriction in the field of present perfect use through processes of discrimination. As it is difficult to attribute Sandrine's sudden adoption of target language rules to recent exposure to positive input, it is more feasible to suppose that she is attempting to bring her past tense output into line with pre-stored declarative knowledge. This knowledge, acquired in formal language learning settings, has remained inert until she has had sufficient exposure to the language to be able to use the knowledge to generate production rules corresponding to target language norms.

### *3.1.3. Other examples*

The gradual progression of Sandrine's use of the past tense from productions based on distinctions operating in her first language to productions conforming to target language usage can be traced in the following table, which presents every occurrence of past tense use in Sandrine's second language output over the six tasks she performed.

The progression over the six tasks shown in table 2 indicates a shift and reduction of the functional field of application of the present perfect and the establishment of the simple past as the main past tense. In the light of this progression, utterances that conform to both first and second language rules are assumed more likely to be based on first language rules in occurrences 1–22 and on second language rules from occurrence 23 onwards. Apart from the cases where the verb form is defective (8, 11, 12, 13, 16 and 31), the cases of present tense use (15 and 19) and the one case already mentioned where the present perfect rule is wrongly applied (22), only utterances 1 and 10 do not conform to the general trend. Rather than constituting instances of application of the second language rule, these utterances seem however more likely to be based on other competing rules: in occurrence 1, Sandrine's utterance ("yes I did, I got irritated") may be input-driven, resulting from a simple repetition of the simple past form in the question that her partner reads from the questionnaire used in the discussion task; as for the present perfect form in occurrence 10 ("how long have you been in prison"), it may constitute a memorised chunk or be following an alternative rule associating present perfect use with the interrogative pronoun 'how long'.

In the framework of this progression, the cluster of defective forms appearing at the intermediate stage can be considered to be a product of the

Table 2. All cases of Sandrine's past tense use over the six tasks

Task	Occurrence	Utterance	Rule
1	1	yes I did, I got irritated	L2
2	2	I studied psychology and sociology during my first year	L 1 / L 2
2	3	I have had a degree two years ago	L1
2	4	I worked before	L1
2	5	I have lost my parents	L1
2	6	when I was children	L1/L2
2	7	I have eh <b>j'étais</b>	L1
4	8	what was you doing to be in prison	defective form
4	9	you have earned money illegally	L1/L2
4	10	how long have you been in prison	L2
4	11	you are lost the liberty	defective form
4	12	when are you passed in court	defective form
4	13	people wasn't confidence in you	defective form
4	14	I have passed a casting in France TF1	L1
4	15	the casting is good	present tense
4	16	I was worked at TF1	defective form
5	17	since the creation of the organisation, states have been controlled	L 1 / L 2
6	18	I learned English	L1/L2
6	19	at the same time I look after children	present tense
6	20	in the family where I lived	L1/L2
6	21	I looked after children	L1/L2
6	22	they have been	defective L2 present perfect rule
6	23	I visited Canterbury	L2
6	24	I saw Westminster	L2
6	25	I went in London	L2
6	26	I visited some monuments	L2
6	27	I lived in Slough	L2/L1
6	28	I went in Scotland	L2
6	29	I visited Loch Ness	L2
6	30	I didn't see the monster	L2
6	31	where were you live in USA	defective form
6	32	why have you been in USA	L2



same combination of metalinguistic reflection and first language mediation, and thus to fit in to a coherent pattern of linguistic behaviour and development.

Three of these defective forms seem to be attempts to use the past progressive tense: “*what was you doing in prison*” (8); “*I was worked at TF1*” (16); “*Where were you live in the USA*” (31), and indeed this tense is appropriate in all three contexts. The appearance of this third past tense functional category in Sandrine’s output establishes a discrimination with regard to the simple past and the present perfect and is likely to have repercussions on her representation and use of these other two tenses.

Finally, example 11 presents two cases of the same hybrid form: “*you are lost*” and “*when are you passed*”, a mixture of the present perfect (expressing past time) and the present continuous (expressing future time) which seems to provoke a certain misunderstanding:

- (11) H It’s very difficult for me to: to live in that prison, I don’t I don’t like this.  
 S *you are lost* eh the free **la liberté** the liberty  
 H yes  
 S and eh: it’s eh OO and eh *when are you passed* in court court of appeal **passer devant je ne sais pas**  
 H **il a déjà fait**  
 S **non il revient. Il y revient.**  
 H I I don’t know the exact date.

Whatever production rule underlies Sandrine’s utterance, a problem of form gives rise to reflection on function in the ensuing exchange in the first language where the speakers disagree about whether the appearance in “court of appeal” is in the past or in the future<sup>7</sup>: “**il a déjà fait**” → “**non il revient**”. Despite the incorrectness of the verb forms, the underlying cognitive processes which are set off can, in the theoretical perspective adopted in this study, be regarded as having a role in the development of Sandrine’s metalinguistic awareness concerning form-function mapping related to verb tenses.

### 3.2. Other learners’ development of past tense use

It would seem therefore that the metalinguistic activity underlying Sandrine’s success in the (1997) study can be described not only in terms

of a concern for accuracy as regards target language norms but also as a search for consistency in the production rules she applies, even if these rules are initially modelled on the first language system. In order to explore this hypothesis further, the progression in past tense use over the six tasks of two other students was analysed: Marie<sup>8</sup>, who was categorised as a frequent repairer, and Jean<sup>9</sup>, categorised as an infrequent repairer (Griggs 1997). These two students were chosen because their level of competence in the second language at the beginning of the period of investigation was roughly the same as that of Sandrine.

Table 3. Quantitative analysis of Marie's use of past tenses

Tasks	N° contexts	Correct use	Functional error	Formal error
T1	5	1	4	0
T2	7	0	7	0
T3	0	0	0	0
T4	3	2	0	1
T5	1	0	1	0
T6	12	7	4	1

Table 3 indicates that Marie's output is characterised by the attention she pays to formal accuracy when using verbs: she makes only two formal errors in 28 contexts. On the other hand, in all cases but one, the verb forms she employs in the first two tasks are functionally inappropriate to the context. A closer analysis of her utterances in table 4 reveals that during these two tasks she only once (occurrence 2) uses the simple past, which is the appropriate tense, and that in the rest of the 12 contexts, apart from occurrences 5 and 8, she employs the simple present. Her output in these first two tasks can therefore be considered to be based on the production rule: "If you want to express past time then use the present simple". As such, what appears to be a reduction strategy (Faerch & Kasper 1983) can also be considered as a way of achieving a certain functional regularity in the use of verbs.

The present perfect appears for the first time, although functionally inappropriate, in occurrence 8, accompanied by a metalinguistic comment in French in which Marie insists that the form is correct: "**ça se dit**" ('you can say that'). Later on, during tasks 4 and 5, the present perfect emerges as the sole past tense in occurrences 14, 15 and 16, and in 13 if "you are commit" is to be considered as an attempt to produce a present perfect

Table 4. All cases of Marie's past tense use over the six tasks

Task	Occurrence	Utterance
1	1	in school eh I not eh I am not good in <b>langues</b>
1	2	I wanted to learn Spanish
1	3	in grammar test I'm not good
1	4	I am good in English
1	5	after it would become too complicate
2	6	I work in a hospital for 6 months
2	7	I can I can work with many old age
2	8	eh I have had <b>ça se dit</b> have had good contacts with them
2	9	in addition I work with children
2	10	I can propose them with many activities
2	11	she speak speaks about it
2	12	she tell, told, tell, tells me, she tells me
4	13	but you are commit bad things
4	14	you has no no you have robbed
4	15	you have you have robbed money
5	16	they have done a bad education at their children
6	17	I went to England last year
6	18	I would like to pass the certificate
6	19	I study one year in Cambridge
6	20	the week I live I lived in Cambridge
6	21	I worked in MacDonal'd's
6	22	I meet other French
6	23	who I visited in London
6	24	I kept the contact with the friend
6	25	who study with me and with friends
6	26	who work with me
6	27	you are you have spent one year in England
6	28	where did you go in the USA

form, and, like the present simple previously, it seems to be used indiscriminately to express past time.

Finally, in task 6, the simple past, which appears only for the second time, is adopted by Marie as the main past tense, and is used 6 times in the

12 possible contexts. The present simple still appears 4 times, and the present perfect once (occurrence 27), its use being on this occasion appropriate to the context. However, despite the consistent pattern in her use of tenses, Marie, unlike Sandrine, does not yet display an ability to make a functional distinction between the past simple and the present perfect.

Table 5. Quantitative analysis of Jean's use of past tenses

Tasks	N° contexts	Correct use	Functional error	Formal error
T1	5	2	2	1
T2	4	1	0	3
T3	0	0	0	0
T4	4	1	2	1
T5	0	0	0	0
T6	10	3	4	3

Table 5 indicates that Jean pays less attention than Marie to producing correct target language forms, making 8 formal errors in 23 contexts. Neither does he adopt any regular pattern in his use of tenses: table 6 shows that in the first 5 tasks all three tenses occur haphazardly, the present perfect 6 times, the simple present 4 times and the simple past 3 times, and their use does not seem to be related to any underlying functional rule. Task 6, which generates only 3 cases of correct past tense use, reproduces a similar distribution of verb forms to that of the previous tasks – 4 present perfect forms, 4 simple present forms and 2 simple past forms – without there being any evidence still that a functional distinction is being made. So what characterizes Jean's output, in contrast to that of the two other more successful learners, is that it reveals both less attention to formal accuracy and very little consistency in the past tense rules he applies.

#### 4. Conclusion

Unlike a traditional longitudinal analysis, where second language acquisition is envisaged as a gradual insertion into the prefabricated mould of the target language, the cognitive perspective adopted in this study has described interlanguage development in terms of the rules and strategies the learners themselves seem to create in order to achieve communication goals in the

Table 6. All cases of Jean's past tense use over the six tasks

Task	Occurrence	Utterance
1	1	but the exercise we have do we have did
1	2	some ones I am right
1	3	some others I have wrong
1	4	when I was in primary school
1	5	I had a good result
2	6	So you have not repeat your year
2	7	I was get fired
2	8	he has no resisted
2	9	have you made have you made your service
4	10	all the men have made a tes before shoot the film
4	11	when I begin the first films
4	12	I enjoy but now I am boring
4	13	Have you already got some relations with a Japanese
6	14	I have no time
6	15	I work all the time
6	16	I work in a bar
6	17	I was a barman
6	18	it was very <b>berk</b>
6	19	I sleep a lot
6	20	you have just spent a year in the USA
6	21	how many have you visit
6	22	have you visit <b>Floride</b>
6	23	have you visit Hollywood

second language. Interlanguage development is thus regarded not as a sequence predetermined by universal grammar but as a process of construction based on an interaction between internal cognitive processes and target language input. In such a perspective the main contribution of communication tasks to the development of second language production skills would seem to be the interactional setting they provide for communicative language practice. The effectiveness of communication tasks resides not so much in the natural communicative and acquisitional processes they trigger as in the way learners use them to construct production rules appropriate to

communicative needs while adjusting these rules to target language requirements. In this framework, controlled and collaborative metalinguistic activity has been shown to play a central role both in building second language discourse and in accompanying transitional phases of second language development. The main hypothesis that has been formulated in this study, and that has been backed up by some empirical evidence, is that, for successful learners, metalinguistic activity is reflected both in the attention they give to formal accuracy and in their application of a consistent pattern of functional rules which may or may not conform to target language norms.

### Transcription rules

pause	0
filled pause	eh
syllable lengthening	:
rising intonation	,
falling intonation	.
overlaps	+...+ +...+
first language items in bold print	<b>matière je ne sais pas le dire</b>
language structures focussed on in the study in italics	<i>I studied psychology</i>

## Notes

1. A common typology of communication tasks, for instance, is based on three types of gap (see Nunan 1989): 1) *information gap activities*, which require the transfer of given information from one person to another or from one form to another; 2) *opinion gap activities*, which involve identifying and articulating a personal preference, feeling or attitude in response to a given situation; 3) *reasoning gap activities*, where new information is derived from given information through processes of inference, deduction, practical reasoning etc.
2. The first group, of a lower intermediate level, were given four types of task to perform: 1) a spot the difference task where the learners had to find the twelve differences between two almost identical pictures by describing their own picture and not looking at their partner's; 2) a simulation activity where the actors, survivors from an accident in the middle of a desert, had to choose the objects they needed to survive; 3) role plays based on interview situations; 4) discussions based on questionnaires. The second group, of a slightly higher level, performed the last two types of task only.
3. See transcription rules at the end of the text.
4. According to Vygotsky (1962: 109) the first language functions as an instrument of mediation and control in second language development. This postulate corresponds also to Weinreich's (1953: 10) concept of *subordinative bilingualism* according to which "the referents of the signs of a language in the process of being learnt can be not real things but equivalent signs in a language already known". Weinreich's concept has been taken up more recently by researchers working in the perspective of Levelt's (1989) language production model. Kroll (1993), for example, envisages second language as a progression from *subordinative bilingualism*, where connections between the lexical representations in the two languages predominate, towards *compound bilingualism*, where the two respective lexical stores are based on a common conceptual representation.
5. The name is fictional to preserve anonymity.
6. Apart from these fortnightly, sixty-minute oral sessions in the language laboratory, Sandrine, whose main subjects were economics and law, had little exposure to the target language, the contact being limited to one hour a week of work on specialised texts in English and to one hour every two weeks devoted to oral comprehension activities. It is also unlikely she spent much of her free time studying English on her own, as her university work load was extremely heavy. Consequently, any progress made over the period covered by the study can be attributed to a large extent to work realised during and immediately after the performance of these communication tasks.
7. This task is a simulated interview with the French business tycoon Bernard Tapie after his imprisonment on corruption charges.
8. See note 5.
9. See note 5.

## References

- Anderson, John  
1982 Acquisition of Cognitive Skill. *Psychological Review* 84 (4): 369–406.  
1983 *The Architecture of Cognition*. Cambridge, Mass.: Harvard University Press.
- Bange, Pierre, Carol Rita and Peter Griggs  
2002 Genèse du système d'opposition présent/passé (sous-système du parfait) en allemand chez une apprenante francophone en immersion. *Marges Linguistiques* 4.
- Bialystok, Ellen  
1990 *Communicative Strategies: A Psychological Analysis of Second Language Use*. Oxford: Blackwell.
- Bialystok, Ellen and Michael Sharwood Smith  
1985 Interlanguage is not a state of mind: An evaluation of the construct for second language acquisition. *Applied Linguistics* 6: 101–17.
- Duff, Patricia  
1986 Another look at interlanguage talk: Taking task to task. In *Talking to learn*, Richard Day (ed.), Cambridge: Newbury House.
- Ellis, Rod  
1990 *Instructed Second Language Acquisition*. Cambridge: Blackwell.  
1994 A Theory of Instructed Second Language Acquisition. In *Implicit and Explicit Learning of Languages*, Nick Ellis (ed.), London: Academic Press.
- Faerch, Claus and Gabriele Kasper  
1983 *Strategies in Interlanguage Communication*. London: Longman.
- Griggs, Peter  
1997 Metalinguistic work and the development of language use in communicative pairwork activities involving second language learners. *Views on the Acquisition and the Use of Second Languages*. Barcelona, EUROSLA 7 Proceedings: 403–416.  
2000 Rôle des stratégies d'apprenants dans des tâches communicatives réalisées en binômes. *La didactique des langues dans l'espace francophone: Unité et diversité*. Grenoble, Actes du 6<sup>e</sup> Colloque International ACEDLE: 161–170.
- Karmiloff-Smith, Annette  
1992 *Beyond Modularity. A development Perspective on Cognitive Science*. Cambridge, Mass.: MIT Press.
- Krashen, Stephen  
1982 *Principles and Practice in Second language Acquisition*. Oxford: Pergamon.
- Levelt, Willem  
1989 *Speaking. From Intention to Articulation*. Cambridge MA: MIT Press.



- Kroll, Judith  
1993 Assessing conceptual representations for words in a second language. In *The Bilingual Lexicon*, Robert Schreuder and Bert Weltens (eds.), Amsterdam: John Benjamins.
- Long, Michael  
1983 Native speaker/non-native speaker conversation and the negotiation of meaning. *Applied Linguistics* 4: 126–141.
- Long, Michael and Patricia Porter  
1985 Group work, interlanguage talk and second language acquisition. *TESOL Quarterly* 19: 207–225.
- Mclaughlin, Barry  
1987 *Theories of Second Language Learning*. London: Edward Arnold.
- Nunan, David  
1989 *Designing tasks for the Communicative Classroom*. Cambridge: Cambridge University Press.
- Nussbaum, Luci, Amparo Tuson and Virginia Unamuno  
2002 Procédures de contournement des difficultés de langues dans l'interaction entre apprenants. In *Discours, action et appropriation des langues*, Francine Cicurel and Daniel Véronique (eds.). Paris: Presses Sorbonne Nouvelle.
- Pica, Teresa  
1987 Second language acquisition, social interaction and the classroom. *Applied Linguistics* 8: 3–19.  
1994 Research on negotiation: what does it reveal about second language learning conditions, processes and outcomes? *Language Learning* 44: 493–527.
- Pica, Teresa, L. Hollyday, N. Lewis and L. Morgenthaler  
1989 Comprehensible output as an outcome of linguistic demands on the learner. *Studies in Second Language Acquisition* 11: 63–90.
- Sharwood-Smith, Michael  
1993 Input enhancement in instructed SLA: theoretical bases. *Studies in Second Language Acquisition* 15: 165–179.
- Schmidt, Richard  
1990 The role of consciousness in second language learning. *Applied linguistics* 11: 129–158.  
2001 Attention. In *Cognition and Second Language Instruction*, Peter Robinson (ed.), Cambridge: Cambridge University Press.
- Schneider, Walter and Richard Shiffrin  
1977 Controlled and automatic information processing. *Psychological Review* 84: 1–57.
- Skehan, Peter  
1996 A framework for the implementation of task-based instruction. *Applied Linguistics* 17: 38–62.  
1998 *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.

Swain, Merrill

- 1985 Communicative competence: some roles of comprehensible input and comprehensible output in its development. In *Input in Second Language Acquisition*, Susan Gass and Caroline Madden (eds), Rowley, Mass.: Newbury House.

Swain, Merrill and Sharon Lapkin

- 1995 Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics* 16: 371–391.

Tomlin, Russell and Victor Villa

- 1994 Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition* 16: 183–201.

Varonis, Evangeline and Susan Gass

- 1985 Non-native/native conversations: a model for negotiation of meaning. *Applied Linguistics* 6: 71–81.

Vygotsky, Lev

- 1962 *Thought and Language*. Mass.: MIT press.

Weinreich, Ulrich

- 1953 Reprint *Languages in Contact: Findings and Problems*. New York: Linguistic Circle of New York. Original edition by Mouton, The Hague, 1974

Zobl, Helmut

- 1995 The 'acquisition-learning' distinction. *Applied Linguistics* 16: 35–56.

# Language learning in content-based instruction

*Heini-Marja Järvinen*

Content-based language instruction is a relative newcomer in the European educational context. This paper presents some more or less established forms of content-based instruction focussing on Finnish implementations of content and language integrated learning (CLIL). The paper also provides a discussion of models and theories of language learning that are compatible with language learning in content-based instruction, and suggests a model for the theoretical framework of the study presented in chapter 4. In the study, the syntactic development of relativization of 7 to 11 year-old Finnish CLIL students was tracked as measured by elicited imitations and grammaticality judgments. The paper concludes with a discussion of the results of the study, their implications as well as some problems and questions related to the study and content-based instruction in general.

## 1. Introduction

While learning in a non-native language is an old phenomenon, content-based language instruction is a relative newcomer in the European educational context. Its increasing popularity can be tracked down to societal and political factors on the one hand and to language learning research and language teaching methodology on the other.

At the beginning of a new millennium, the importance of language skills is generally acknowledged in Europe, where the forming of the European Union has led to pan European integration of language teaching. So far, the results of this large-scale work have been materialized in the publication of two influential documents. In *White Paper on Education and Training. Teaching and Learning – Towards the Learning Society* (1995), communicative proficiency in two non-native languages were set as a goal for all language learners in the European Community. The *Common European Framework* (2001) has been compiled to facilitate mobility within the Union by integrating language teaching and learning in Europe.

Not only societal factors but also advancements in SLA research and language teaching have contributed to the spread of content-based language

teaching. Research on Canadian immersion programs, which has been carried out consistently from the 1960's onwards, has produced rich and reliable data on language learning and teaching in content-based instruction. In the domain of language teaching methodology, the communicative approach is based on the use of language in realistic situations, thus meaning, i.e. functional language is emphasized in relation to form.

Chapter 2 presents some more or less established forms of content-based instruction, discusses significant features and outcomes of content-based instruction. Then, Finnish implementations of content and language integrated learning (CLIL) are addressed. The aim of chapter 3 is to provide a selected overview of models and theories of language learning that are compatible with language learning in content-based instruction. Specifically, it comprises the theoretical framework of the study presented in chapter 4.

## **2. Content-based language teaching**

### *Terms*

Content-based instruction encompasses a wide range of models and implementations, such as immersion, language-enhanced content learning, mainstream bilingual education, plurilingual education, two-way bilingual education, and content-based language teaching. The terms relate to 'teaching methods' with their own typical features, commonly reflecting different emphases on the language and content components. In Europe, a generic term has been introduced (Marsh & Marsland 1999) to encompass all forms of content-based instruction: CLIL for "content and language integrated learning". In this chapter, content-based instruction (CBI) is used as a generic term to refer to all forms of instruction in which content and language are combined. Content-based language teaching (CBLT) is used to refer to content-based instruction whenever language teaching is emphasized. Finally, content and language integrated learning (CLIL) is used to refer to the recent European implementations of CBI, the present study included.

### *Current forms of CBI*

Brinton, Snow & Wesche (1989: vii) define content-based instruction as "the integration of content learning with language teaching aims. More specifically, it refers to the concurrent study of language and subject matter, with the form and sequence of language presentation dictated by content" (Brinton, Snow, and Wesche 1989: vii). A literary interpretation of the defi-

tion yields a “strong” form of content-based instruction (Wesche & Skehan 2002), in which content alone determines what language is used in teaching content. A “weaker” version would imply a stronger emphasis on the language component and the language learner, and less emphasis on content (Wesche & Skehan 2002). Obviously, this is a continuum incorporating a variety of versions of content-based instruction with different emphases on content and language. In fact, it is difficult to imagine a strong version of content-based instruction with no, even implicit, adjustments to the level of language of instruction to make teaching comprehensible for second language learners.

CLIL refers to “initiatives, in which the learning of second/foreign languages and other subjects has a joint curricular role in education.” (Marsh and Marsland 1999: 9). The definition of CLIL emphasises the equally important role of both content and language (Marsh and Marsland 1999: 9). “A joint curricular role” of content and language would ideally envision a systematic integration of both content and language syllabi in accordance with the national curricula. This is a desirable goal, but its attainment will require a lot of practical work, such as collaboration between content and language teachers, in-service teacher education and materials design. On the other hand, “a joint curricular role” may be interpreted more loosely tied to subject-specific curricula. This seems to be the case at the time of writing, judging by the national inventories made at the end of 1990s (Nikula and Marsh 1996, 1997).

The majority of content-based models (Brinton, Snow and Wesche 1989) cater for the language learner in various ways. *Theme-based* versions refer to content-based instruction in which the language syllabus has been arranged around the core thematic syllabus. If the language component is left intact, theme-based instruction represents a stronger form of content-based instruction, but more often than not at least in the Finnish context, the choice of themes is made with the language and the language learner in mind. Theme-based forms of content-based instruction are particularly useful at primary level, as they form part of an otherwise entirely or mostly theme-based curriculum, of which some themes are selected for teaching in a non-native language. Another format of content-based instruction, *adjunct* courses, are content-based courses in which separate but concurrent language courses are provided which support the second language learner in learning content-related language. Content instruction is directed at native and non-native learners, but only the latter participate in language teaching supporting content instruction. Adjunct courses are common in a second-language study context, e.g. in the U.S, but they are not common in the

Finnish content-based instruction. However, the basic idea of teaching content-related language in contingent foreign language lessons could be adapted in content-based subject teaching at secondary level. The third form of content-based instruction; the *sheltered* format, recruits students with identical linguistic background and teaches them content in the language they are learning. In Finland, the majority of forms of content-based instruction share features of the sheltered format: learners speak Finnish as their L1, and they are more or less at the same level in their foreign language skills on entrance to the content-based program. In particular, this concerns those forms of content-based instruction that use selection criteria in accepting students. According to a survey of teaching content in a foreign language in Finland (Nikula and Marsh 1997), the majority of content-based programs do not apply any selection criteria. Most primary content-based groups consist of intact classes; at upper secondary level students may opt for courses taught in a foreign language, and, if necessary, selection is made on the basis of students' recent grade in the language of instruction.

### *Immersion*

One of the most successful forms of sheltered content-based instruction is immersion. Of all forms of content-based instruction, immersion programs have probably influenced content-based instruction the most. Language learning in immersion has been amply, rigorously and systematically researched from the outset of immersion programs in Canada from mid 60s and onwards (Stern 1985; Cummins & Swain 1986; Wesche 2001). The results have been consistently good in terms of both language and content learning. The best-known form of immersion instruction, *early immersion*, has yielded the best results (see e.g. Genesee 1987; Lapkin et al. 1990): children, who have been totally "immersed" in the immersion language for one or two years in kindergarten and/or at school and have later been taught at least half of the academic curriculum in the immersion language and half in their first language, reach functional skills in the immersion language without any harm done to their first language or academic achievement in general. In addition to purely linguistic gains, the learners develop more positive attitudes towards the immersion language, its speakers and its culture, and they are more open and natural in spoken expression in particular. Other forms of immersion, such as *partial immersion* (less than half of the curriculum taught in the immersion language) and *late immersion* (start in grades 3, 5 or even 7) have also led to language learning results that outperform those attained in traditional language teaching (Wesche 2002).

By definition, immersion programs share the general features of content-based instruction (Wesche and Skehan 2002), of which the most profound one is the assumption that learners learn both content and language at the same time. In teaching academic content, expository text type is frequent. Thus, instruction in compatible literacy skills, reading strategy instruction included, is likely to facilitate comprehension and in-depth learning of subject matter (Mohan 1986) after the first two years into immersion, when spoken skills are no longer the only goal of language instruction. Language input and interaction are adapted to accommodate the learners' language level, not only in terms of comprehensibility but also in terms of optimal language development. The two latter features are characteristic of the sheltered content-based instruction format: learners with homogeneous language and cultural background are enrolled in immersion programs, the academic curriculum is taught in a second language, whose level is adapted to accommodate the language learners' level and learning needs.

Moreover, immersion programs incorporate pedagogical features, operationalized from a number of assumptions about language learning (Wesche 2002). These are an early start in the immersion program; beginning with total immersion in the second language, followed by a gradual decrease of the amount of the second language and simultaneous increase of the amount of the first language over several years, and ultimately stabilising at some predetermined ratio, commonly 50-50, for a number of years. The goal of immersion language learning is functional bilingualism, which is regarded as best attained by native speaker or bilingual teachers, and contextualised and motivating language teaching. Until recently, there has been very little explicit teaching of the immersion language. However, research into language learning in immersion showed that immersion students' production and accuracy were considerably inferior to their native-level comprehension skills. As a result, more attention to language form and increased opportunity for demanding output are considered necessary for continued language learning in immersion (Swain 1993).

#### *Context of the present study*

The context of the study reported on below is CLIL in Finland, a country in northern Europe with some five million inhabitants and two official national languages, Finnish and Swedish. Finns have always valued knowledge of many foreign languages for at least one obvious reason: Five million Finns speak Finnish, a Finno-Ugric language which is very different from the major languages of trade and communication spoken by hundreds of

millions of people in the world. This is one reason for the surge in popularity of content-based instruction. Another is the national decision made at the beginning of the 1990s, which was to allow instruction in a non-native language (i.e. other than Finnish, Swedish, Lappish or the Romani language) in general education. In addition to this initial incentive, various integration processes in Europe and encouraging results from language teaching methods, communicative language teaching in particular, have contributed to the spread of CLIL to all levels of education. An increasing number of languages, such as Swedish, French, German and Russian, are increasingly used as a medium of instruction in addition to English, which is by far the most popular one. Figure 1 presents some of the characteristics of models currently implemented in CLIL in Finland.

---

#### LANGUAGES

---

- The majority of the instruction takes place in the students' L1, Finnish.
- Finnish is consequently used in the teaching of literature and mother tongue.
- The amount of class time devoted to teaching in English varies greatly.
- In CLIL, the “norm” is no less than 20 percent of class time taught in English.

---

#### TEACHERS

---

- Finnish teachers, who are non-native but relatively fluent speakers of English, teach in Finnish and in English, (native speakers are occasionally available as language models)
- Primary level teachers commonly teach most subjects on the curriculum.
- Secondary and upper secondary level teachers are subject teachers with adequate English skills to teach their subject in English, but usually no formal studies in the non-native language of instruction.

---

#### SUBJECTS

---

- A wide range of subjects is taught in a non-native language
- The tendency seems to be that subjects with concrete (contextualised) content and relatively little cognitive load to the learner are favoured at elementary level of comprehensive school (environmental studies, music, arts, crafts, and PE); and
- at lower secondary level (home economics, biology, geography, and history).
- In upper secondary school, the most popular subjects are history and geography, but the range of subjects taught in CLIL is considerable more variable than at lower levels.

---

#### LEARNERS

---

- all ages: K – 18+
  - all levels of education: kindergarten, general education, vocational sector, university
  - intact classes, optional courses
- 

(Fruhauf et al. 1996; Nikula & Marsh 1996, 1997).

*Figure 1.* Features of implementations of content-based instruction (CLIL) in Finland



### 3. Language learning in content-based instruction

It has taken and it will take SLA research a lot of time and effort to explore the extremely complex phenomenon of second language learning. Obviously, there are many factors that are involved in language learning in immersion and other content-based instruction (e.g. Grabe and Stoller 1997: 5–21) that are involved in language learning in these programs, and it can only be conjectured that the factors that contribute to the successful language-learning outcomes are the ones that make language learning different from “traditional” language teaching. The most obvious of them, the relationship between meaning and language, enhanced opportunity to learn, and time-on-task, will be discussed in this chapter.

In the majority of content-based models, language learning is a by-product to content learning and the language syllabus is derived from the content syllabus. In this sense content-based language learning is “meaning-based” and probably more motivating to learners as it combines the two goals of academic subject learning and language learning. Language is a tool of learning relevant academic content; and as such its use in the classroom is real and thus potentially more challenging, motivating and more pushing (Swain 1993) than for example in communicative language teaching, which is more “meaning-oriented” in the sense that the practice of language functions and situations are classroom role-plays or simulations of authentic language use.

In addition to focus on meaning, and maybe even more important in content-based instruction than has been acknowledged (Takala 1996: 9), is availability of more time on language learning, or more *opportunity to learn*. Time for learning is greatly increased when the duration of content-based instruction is years rather than months or weeks and when content-teaching time is also language-teaching time. It is commonly agreed that extended exposure to rich and meaningful language input (*comprehensible input*, Krashen 1985) is adequate for up to native-like comprehension skills to develop, as in immersion programs, but it is not enough for attainment of accuracy and native-level production skills. In addition to accurate and idiomatic production skills, academic study requires the learner to learn content-related language skills, i.e. to achieve *cognitive academic language proficiency* (CALP, Cummins 1984). According to Cummins, immersion students learn everyday conversational fluency (*basic interpersonal communication skills*, BICS) in a matter of two years, but attainment of academic language proficiency takes five to six years in intensive immersion, and considerably longer in less intensive content-based instruction. To sum up, content-based language instruction provides more opportunity to learn

than language teaching alone, but access to comprehensible input and processing for meaning can only take the second-language learner up to a certain level of language proficiency, whereas native-like proficiency requires some form of attention on language form (Long and Robinson 1998, Doughty 2001). The assumption is that the more advanced the learners are, the more rigorous the intervention must be to prevent fossilization (*comprehensible output hypothesis*, Swain 1985).

Content-based instruction, which provides an extended exposure to rich and complex input, is a potential candidate for implicit learning, which is commonly defined as *a process whereby complex knowledge is acquired about a rich and complex stimulus environment. It is a process which takes place "automatically, naturally, simply and without conscious operations, simply as a result of experience of examples"* (N. Ellis 1994: 1; 1995: 123). In other words, implicit processes identify and extract complex regularities from the stimulus environment. The regularities are then stored as memory representations (schemata) and may be later used as "rules" in instances that resemble the original examples. All this takes place unconsciously, implicitly, without the learner being aware of the process.

The *connectionist* view assumes that the human memory is capable of building prototypical representations on the basis of emerging regularities in the vast number of associations in input and already existing representational associations. Thus, connectionist learning is content- and structure-sensitive; the emergent generalized and transferable structure is ultimately drawn on active competition between the numerous available cues in the input (Ellis 2001: 66–67). Connectionism posits no role for innate predetermined linguistic universals, which are the foundation of generativist accounts, deriving from Chomsky's *Universal Grammar* theories (The Minimalist Program being the latest version, Chomsky 1996). Universal Grammar is a theory of a restricted number of universal principles and associated parameters with two (or more) parameter values each. Parameter values are triggered by salient cues in the linguistic environment. In learning a second language, parameter-resetting may be a cumbersome process, involving the changing of already existing L1 settings for new values on the basis of second language input. According to UG, successful completion of parameter-setting is obvious in a sudden and complete emergence of new parameter values. The role of Universal Grammar in second language acquisition is highly disputed. Some theoreticians, such as Gregg (2001), view Universal Grammar or a Universal Grammar type of property theory as a necessary component of a SLA theory. Property theories are needed to explain *why* learners acquire the knowledge of an L2, whereas transition theories explain *how*

learners come to acquire the knowledge of an L2. Others, such as cognitivists, deny the role of innate processes and view language learning as identical to any learning emanating from general inductive processes (Ellis 2001: 38).

The innatist and connectionist views assume that learning takes place implicitly. The role of focused attention is minimal or at least indirect in these processes. However, cognitive and constructivist notions of language learning view attention as a fundamental component in any language learning (Ellis 2001; Schmidt 1990, 2001). The role of implicit processes is minimal. Schmidt suggests that previously learnt material in LTM can be activated by unattended incoming stimuli, but there is no evidence of learning of anything new (Schmidt 2001: 25–26, 31). In particular, this applies to unconscious registration of L2 syntactic structures previously unencountered in L2 input but stored in long term memory as representations of corresponding L1 syntactic categories.

In the course of memory research, the role of *working memory* (WM) has shown to be central in language learning (Ellis 2001: 33–34). Working memory is a processing system with limited resources and access to *long term memory* (LTM) and to external input. WM consists of three components. The most important of them is the *central executive*, or *supervisory attentional system* (SAS), which controls all kinds of informational flow in WM; activates and inhibits processes and resolves conflicts and allocates attentional resources. In dealing with the limited attentional resources that are available, the central executive is aided by *two short term memory* (STM) slave systems, the *phonological loop* for verbal information and the *visuo-spatial sketchpad* for visual and spatial information. These two slave systems are limited-capacity systems in that the former is able to deal with a restricted amount of serial verbal material and the latter with one picture at a time. To date, little is known about what exactly takes place in WM. It is assumed, however, that there are cognitive microprocesses with a momentary action span, such as *selective attention* to external stimuli, *cognitive comparison* of stimuli in WM with activated material in LTM for noticing gaps, mismatches etc. Another type of processes are continual, more or less automatic macroprocesses, such as *internationalisation of input*, *mapping* of forms to meanings and grammatical functions, *analysis* and *restructuring*. Internationalisation of input is essentially language learning or interlanguage development. Restructuring may be quantitative (e.g. transformation of declarative knowledge into procedural knowledge) or qualitative (e.g. *automatization* of procedural, already restructured material (Doughty 2001). The attainment of automaticity and fluency (DeKeyser 2001) is

dependent on *chunking*. Chunking is building up larger memory units by creating associations between smaller units. The result is a permanent set of associations in LTM storage, but chunking is possible at all levels of representation. Chunking is recursive, building up hierarchies in LTM storage, thus providing structure and organization in memory (Ellis 2001).

There is evidence that there are individual differences in WM capacity. Assuming that selective attention, i.e. noticing (Schmidt 2001), is crucial to language learning, learners with more WM capacity are potentially better language learners than learners with less WM capacity. However, it may be that WM capacity is crucial to learning only when it is challenged (Sawyer and Ranta 2001, 342–343). There is also evidence that the STM span for encoded material is longer than the originally thought five seconds. It is possible to hold encoded material in STM store for up to 40 seconds, and the capacity of the memory span is greatly enhanced if the items processed are already stored in LTM and activated (Doughty 2001: 226).

#### 4. The study

The purpose of the study was to find out how the acquisition of English in a content-based (CLIL) program progresses from the first grade through to the fifth grade as reflected by the syntactic development of relativization (see below). Elicited imitations and grammaticality judgments were used as a measure of interlanguage development of the target feature. For comparison, the imitations and judgments were administered to CLIL students and to peer controls in mainstream language teaching.

##### *Relativization*

Successful imitation of relative clauses requires that the learner be capable of simultaneous processing of a number of separate items which together form the relative clause within a matrix clause. In terms of the Government Binding account of Universal Grammar, the distance between the components of a relative clause signals the degree of depth of embedding. If the depth is great, more movements are necessary to produce a complete sentence consisting of an embedded relative clause and a matrix clause than with shallow embedding.

The depth of embedding of the elicited test sentences is shown below:

1. The little boy [<sub>CP</sub> that<sub>i</sub> [<sub>IP</sub> t<sub>i</sub> [<sub>VP</sub> likes to play football ]]] lives in the yellow house.
2. The little boy [<sub>CP</sub> that<sub>i</sub> [<sub>IP</sub> you [<sub>VP</sub> see t<sub>i</sub> [<sub>PP</sub> in the picture]]]] is my brother.

3. The baby [<sub>CP</sub> whom<sub>i</sub> [<sub>IP</sub> her mother [<sub>VP</sub> is holding t<sub>i</sub> ]]] is very pretty.
4. The man is looking at the cat [<sub>CP</sub> which<sub>i</sub> [<sub>IP</sub> t<sub>i</sub> [<sub>VP</sub> is sleeping [<sub>PP</sub> on his bed]]]]].
5. The little boy [<sub>CP</sub> whose<sub>i</sub> [<sub>IP</sub> mom [<sub>VP</sub> t<sub>i</sub> is a teacher]]] plays football in a team.
6. The little boy has a new, yellow car [<sub>CP</sub> which<sub>i</sub> [<sub>IP</sub> he likes t<sub>i</sub> [<sub>PP</sub> very much]]].
7. I finally found the book [<sub>CP</sub> that<sub>i</sub> [<sub>IP</sub> I [<sub>VP</sub> had been looking for t<sub>i</sub> .]]]

In terms of the *Accessibility Hierarchy* and the *Relative Clause Parameter* (Berent 1994), the degree of depth is viewed as reflecting the difficulty of acquisition of relative clauses. In terms of memory constraints, stretches of words in which the co-indexed items are distant from each other are difficult to process in short term memory. Similarly, relative clauses where the relative pronoun (,) and its trace (t<sub>i</sub>) are located far from each other are more difficult to process than when the pronoun and its trace are near each other. Thus, some relative clauses are assumed to be easier to acquire and produce than others. Accordingly, sentences 2,3,5 and 7 above would be more difficult to process and acquire than the remaining sentences (1, 4 and 6).

The distance between the relative pronoun and its trace, i.e. the depth of embedding, coincides with the syntactic function of the pronoun in the relative clause. Thus, clauses with subject relative pronouns are easier to process and acquire than object relative pronouns, which in turn are easier to process and acquire than oblique objects and genitives. It is also hypothesized that if the syntactic functions of the matrix clause antecedent and that of the relative pronoun coincide (both are subjects or objects), the processing of the relative clause would be easier than if the functions are different (combinations subject-object and object-subject).

Sentences of considerable length are a challenge to memory processing in STM. If there is no LTM representation for incoming stimuli, their processing in WM is inefficient; slower and more cumbersome than if there is an established chunked representation to be activated to aid WM processing. In the present study, subjects do have an L1 schema of relative clauses stored in LTM. In terms of UG (Hawkins 1989), the sites (NPs) that languages can relativize on are contained in the grammatical component of UG. These sites are then triggered by (the L1) primary language data, which is shown by production of complete relative clauses. In view of the cognitive account (see discussion above, Schmidt 2001), exposure and practice in L1 have led to the chunking of information and storage of relative clause schema in LTM. Thus, if the L2 structure roughly corresponds to the L1 structure, nothing essentially new is learnt, and the L1 relativization schema may even allow implicit processing of second-language relative clause input.

*Subjects*

The subjects were primary school students at the university practice school at the University of Turku. The experimental groups (N=90) consisted of the students in grades 1 through to 5 (groups 1E, 2E, 3E, 4E and 5E). The control groups (N=47) were students from the parallel grades (3C, 4C and 5C).

*Quality and quantity of input*

The main difference between the experimental and the control groups lies in the type of exposure to the English language. The experimental groups were involved in the bilingual stream of the Middle Years Program (MYP), in which part of the instruction is given in English from the first grade onwards. As a rule, a fourth of all instruction (25%) is English-medium teaching. The MYP students participated in code-oriented English classes from the third grade onwards. The control groups were involved in Finnish monolingual instruction in all their content subjects. Their first encounter with the English language at school was in the third grade when they started attending code-oriented, formal English classes twice a week. Hence the lack of controls in the first and second grades.

Table 1 shows the amount of instruction of English as a foreign language and the amount of content and language integrated instruction, i.e. instruction through the medium of English. It should be kept in mind that the data in the table are based on estimates collected from CLIL class teachers. In addition to permanent teaching staff, university practice schools cater for the teaching practice of hundreds of student teachers each year. As a number of student teachers teach CLIL classes, CLIL students are exposed to a wide range of interlanguage inputs, more or less fluent and accurate.

*Table 1.* Amount of formal English instruction (FE) and instruction in English (CBE) (45-minute lessons per week)

	Grade 1		Grade 2		Grade 3		Grade 4		Grade 5	
	FE	CBE	FE	CBE	FE	CBE	FE	CBE	FE	CBE
Control	–	–	2	–	2	–	2	–	2	–
Experimental	–	4–5	–	4–5	2	4–5	2	5–6	2	5–6

In formal English instruction, CLIL students do extensive reading of authentic literature apart from textbook study. In grade 3, they read one book which

contains mostly pictures accompanied by very easy language (no relative clauses). In grades 4 and 5 they read two books each year in addition to grade-level textbooks. If each student has his or her own copy, some reading is done at home, some in class. The teacher chooses the book for each group. To take just a few examples, fifth-graders had no trouble in reading Roald Dahl's *Mathilda*, but the language in *Harry Potter* turned out to be too difficult for most of them. The language level of grade 4 and especially grade 5 literature is relatively authentic and is likely to contain a wide range of language, a number of relative clauses included. It should be noted here that the formal English curriculum does not contain the teaching of relative clauses and accordingly, textbooks do not take up relative clauses in grammar sections. This does not preclude occasional occurrence of relative clauses in texts.

When inquired about the use of classroom English in retrospect, permanent class teachers of CLIL groups did not report on other than occasional use of relative clauses. Oral instructions, questions and explanations consisted usually of formulaic expressions and modified input containing simple main clauses.

#### *Pre-tests*

The English CLIL stream, as well as all Finnish children starting school at the age of seven, are routinely administered a set of tests to find out whether the children are ready to start school. The test battery includes tests of mother tongue and motor skills. As the number of applicants almost every year exceeds the yearly quota of CLIL students, the students are selected on the basis of school entrance tests and teacher judgment.

To ascertain that the experimental and control groups did not differ from each other in any way that could be anticipated in the present study, Raven's progressive matrices test, a test of nonverbal intelligence, was administered to both experimental and control groups. This test has previously been shown to be related to explicit learning of language (Kristiansen 1990). As expected, the only significant differences of means were found between the youngest students (grades 1E and 3C), and the older students, that is, the rest of the subjects. No other statistically significant differences were found.

#### *Tests*

In the present study, *elicited imitation* and *grammaticality judgment* tasks were used to measure the subjects' syntactic development. Elicited imitation tasks were assumed to tap the subjects' implicit language competence and



memory for language. Grammaticality judgments were used to measure the subjects' explicit interlanguage competence, especially the metalinguistic component.

### *Elicited imitations*

Elicited imitation is an elicitation instrument in which the researcher provides an oral stimulus sentence which the subject is required to orally reproduce as accurately as possible. Elicited imitation has long been used to evaluate early L1 acquisition and it has proved to be a valid and reliable method of evaluation in first language acquisition research. Recently, its use has been extended to second language acquisition research (Bley-Vroman and Chaudron 1994: 245–262).

The strongest advantage of elicited imitation is that “it can provide overt, direct evidence of the child’s grammar construction for particularly targeted aspects of grammar, and can do so in highly (linguistically) focused ways” (Lust et al. 1996: 62–63). In terms of second language competence, the advantage of elicited imitation tasks is the ability to measure second language learners’ linguistic competence for material that they have no prior experience of and consequently no LTM representation to aid the chunking of the input. Ellis (2001: 48) gives a number of examples of research on the effect of long-term knowledge on STM for incoming related knowledge, such as words, phonological knowledge, grammatical and semantic knowledge. Like other competence measurement tasks, elicited imitation has the disadvantage of being able to measure competence only indirectly. Elicited imitation behavior may only indirectly relate to the learner’s grammatical competence or syntactic knowledge. One disadvantage is the complexity of the design refinement (see description of the elicited imitation task below). (Lust et al. 1996: 69).

In the elicited production test, the experimenter orally presents, one by one, a series of randomized experimental sentences to the subject who is asked to repeat each sentence as presented. All elicited productions are tape-recorded and the responses are transcribed. At the experimental session, each subject has to demonstrate 100 per cent comprehension of the words to be used. Controlling for the subjects’ knowledge of words used in the testing insures that the results obtained are due to the differences in the structural factors manipulated and not due to lack of knowledge of the L1/L2 lexicon. All the stimulus sentences are equalized in syllable length (15 syllables) and approximately in word length (10 words). The stimulus sentences must be too long to be stored in the short-term memory without



being analyzed (Flynn & Espinal 1985: 100–101; Myles 1995: 249–250), and they need to tax the subject's processing ability just enough so that subjects can and do attempt reconstruction, thus overtly involving their grammar. The grammatical structure of the stimulus sentence is relevant to this processing. (Flynn & Espinal 1985: 100) The model sentences vary only in critical grammatical factors, with all others controlled or held constant (Lust et al. 1996: 59). Sentences are also balanced for semantic content.

### *Grammaticality judgments*

By letting the informants judge the grammaticality of a stimulus sentence, grammaticality judgment tasks allow the researcher to evaluate hypotheses about the informants' grammars fairly directly. Grammaticality judgment tasks unarguably evaluate performance data which in turn give the researcher insight into competence by providing information on allowable and disallowable sentences. Thus, they do not give the researcher direct access to learners' competence; they do, however, offer information about what are possible and impossible sentences in the learner-language (Gass 1994: 303–307). The grammaticality judgment tasks consisted of the following 23 sentences (modified from Gass 1994: 320–321)

1. I heard the girl that the man gave a flower to her.
2. I saw the girl that they boy gave a book to her.
3. We respect the man with whom you danced with him.
4. John admires the woman for whom you wrote the letter.
5. Sam remembers the man about whom they told the story about him.
6. He kissed the woman with whom you were talking with her.
7. I saw the child whose sister ran away.
8. She kissed the girl whose finger hurt.
9. He remembers the man who his brother is a doctor.
10. He likes the girl who her uncle is a baseball player.
11. He greeted the man whom I am smaller than.
12. That's the woman whom I am taller than.
13. He thanked the student whom he is smarter than her.
14. I saw the man who crossed the street.
15. She watched the teacher who was writing on the blackboard.
16. I saw the man who crossed the street.
17. She watched the teacher who was writing on the blackboard.
18. He met the man whom you know so well.
19. She likes the same girl whom I like.
20. I saw the girl that the boy hit her.
21. She surprised the teacher whom the boy thanked her.
22. We washed the baby to whom you had given a doll.
23. He married the woman to whom you wrote the letter.

The subjects' task was to judge whether the sentence was correct or incorrect. The students were provided 45 minutes for the grammaticality judgment task and they were instructed to progress at their own pace. All the participating subjects completed the task within that time.

#### *Method of analysis*

After the elicited imitations had been transcribed, they were coded for subsequent analyses. A matrix was designed in which the raw data consisting of the imitations and grammaticality judgments of each subject were transferred. Each occurrence of an item (word) was coded as 1, if the subject's imitation was a correct imitation of the original stimulus and also another item occurred. If the subject had failed to imitate the particular item, the item was coded as 0 in the matrix.

The data thus achieved were then analyzed. First, the analysis of variance was used to evaluate the differences in means between the experimental and control groups. When statistically significant differences were detected, Tukey's test for multiple comparisons was conducted in order to find out which pairs of groups differed significantly from each other.

## **5. Results**

### *Elicited imitations*

Table 2 presents the means of the elicitations of test sentences by experimental subjects (1E, 2E, 3E, 4E and 5E) and their controls (3C, 4C and 5C).

It appears from the table that the beginners failed to produce any relative clauses at all, whereas the most advanced CLIL students (5E) consistently and successfully imitated 90–100 per cent of all the relative clauses in the stimulus sentences. (Multiplying means by 100 yields the percentages of successful imitations.)

In general, the results indicate that the advanced CLIL groups 3E, 4E and 5E were able to produce imitations of relative clauses at a significantly higher level than the beginning CLIL groups (1E and 2E) and the control groups 3C and 4C (at least at the level of  $p < .05$ , but frequently at the level of  $p < .001$  and  $p < .000$ ). There is a conspicuous difference in means between the two CLIL groups 2E and 3E. The highest control group 5C succeeds statistically significantly better than group 4C on sentence 3 ( $p = .006$ ) and group 3C on sentences 6 ( $p = .039$ ) and 7 ( $p = .042$ ). Other differences between the control groups are not significant.

Table 2. Means of imitations of relative clauses

Grade (N)	Mean1	SD1	Mean 2	SD2	Mean 3*	SD3	Mean 4	SD4	Mean 5	SD5	Mean 6	SD6	Mean 7*	SD7
1E (20)	0,00	0,00	,100	,310	,050	,224	,100	,301	,150	,367	,050	,224	,050	,224
2E (17)	,176	,392	,235	,437	,235	,437	,235	,437	,235	,437	,118	,332	,118	,332
3E (17)	,647	,493	,941	,243	,589	,507	,706	,470	,824	,393	,647	,493	,647	,493
4E (18)	,889	,323	1,00	0,00	--	--	,778	,428	,556	,511	,722	,461	--	--
5E (18)	,944	,236	1,00	0,00	1,00	0,00	,944	,236	,889	,323	,945	,237	1,00	0,00
M <sub>E</sub>	,530	,655	,655	,581	,581	,553	,553	,496	,531	,496	,496	,454	,454	,454
3C (16)	,063	,250	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
4C (18)	,167	,383	,222	,428	,056	,236	,222	,428	,167	,383	,333	,485	,111	,323
5C (13)	,462	,519	,308	,480	,538	,519	,308	,480	,308	,480	,462	,519	,111	,323
M <sub>C</sub>	,231	,177	,177	,198	,198	,315	,315	,265	,288	,265	,265	,074	,074	,074
ALL 137	,416	,495	,482	,502	,345	,477	,416	,495	,394	,490	,409	,493	,328	,471
F	20,5	(7,129)	35,5	22,1	15,0	12,4	14,8	24,1						
Df	(7,129)	(6,112)	(7,129)	(7,129)	(7,129)	(7,129)	(7,129)	(6,112)						
P	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000

Mean 1...Mean 7 = means of imitations of relative clauses for sentences 1 to 7;  
 SD1...SD7 = standard deviations of means of imitations of relative clauses in sentences 1 to 7;  
 ME = Mean for experimental groups (1E, 2E, 3E, 4E, 5E);  
 MC = Mean for control groups (3C, 4C, 5C);  
 F = test value of F-test;  
 df = degrees of freedom;  
 p = p-value

*Grammaticality judgments*

The performances on the elicited imitations and the grammaticality judgments were correlated, and a positive relationship was found between the performance on the false grammaticality judgments and relativization as measured by elicited imitations in the experimental groups ( $r=.62$ ;  $p=.000$ ;  $N=47$ ); that is, those whose imitations of relative clauses were accurate were also capable of identifying incorrect grammaticality judgments with comparable accuracy, and vice versa, those who performed less accurately on elicited imitations also failed to spot false grammaticality judgments. The correlations for performance on elicited imitations and correct grammaticality judgments were not significant in the experimental groups ( $r=.19$ ;  $p=.20$ ;  $N=47$ ). The corresponding correlations for control groups were not significant, but like the experimental, they also performed better on identifying false grammaticality judgments ( $r=.24$ ,  $p=.13$ ;  $N=40$ ) than on identifying correct ones ( $r=.11$ ;  $p=.496$ ;  $N=40$ ).

*Summary of results*

The results of the elicited imitations revealed that the CLIL groups not only produced significantly longer sentences than the control groups but also significantly more complex and more accurate than the control groups. The structures of early imitations of both the CLIL and the control students were in general similar, but a significant increase is apparent between the second and third CLIL group. Where the imitations of 2E are short sentence fragments the imitations of 3E are syntactically fully developed sentences and bear more resemblance to that of the higher CLIL groups, 4E and 5E. The subsequent development, i.e. that between 3E, 4E and 5E is more even, but it is obvious that the performance of 5E is by far the most advanced of all the groups. The most imitations of this group show complete acquisition of relativization. In terms of Universal Grammar, sudden, full-blown emergence of a syntactic structure at some phase of linguistic development is a sign of completed parameter-(re)setting. However, it is not clear how parameter-resetting occurs, how quickly and how it exactly can be identified. In fact, it is not clear if parameter-resetting can be operationalized or whether it is a more metaphorical explanation. It is more likely that the observed interlanguage development is due to years of enhanced input both through subject matter instruction and through language studies. By the end of grade 5, students have done quite a lot of extensive reading (at least five books, some probably more) in English. Thus, they have received a lot of comprehensible input through reading, too. Ample

input and opportunity to process language in the course of months and years have certainly contributed to language acquisition in terms of cognitive skill acquisition. There has been enough time for longitudinally developing processes, such as incremental automatization and chunking. These processes facilitate the limited-capacity processing of the WM and leave space for less automatized processes. It is also possible that the students' short term memory span has been improved as a result of challenging processing of the second language.

## **6. Discussion**

Bilingual streams are highly popular. Many educated parents see them as an opportunity to provide their children better opportunities in the future through bilingual education. In addition, bilingual streams are commonly advertised outside the school district. All this means that some parents are more interested in their children's education than others and they are also willing to devote time and energy to the child's education, for example by participating in school events, taking their children to school by car, planning holiday travels with the child's language development in mind, buying books home or subscribing to magazines in the foreign language, etc. In other words, the parents' values will catch on to the children, who are more motivated to study in a foreign language and more likely to be successful in language learning. Bilingual streams attract more children than can be enrolled. This means that the experimental students were selected students, who participated in a selection procedure prior to entering the bilingual stream at the target school, whereas the controls were local kids entering school in due time. They also were tested for school readiness, but the results could at worst – or at best, and in very few cases – mean delayed school start. This implies that the experimental subjects in the present study may have been an exceptional group, for example in terms of aptitude and motivation. On the other hand, the pre-tests revealed no significant differences in abstract reasoning between the experimental and control peers.

The theoretical framework presented theoretical support and research findings on the significance of STM span and WM capacity in language learning. The empirical results can be viewed as support to these findings. Sawyer and Ranta (2001) suggest further research to confirm and fine-tune the relation between individual STM capacity and language learning. Once this is done, STM measures, such as elicited imitations, may have more applicability in language testing.

The present study corroborates in part the conclusion that CLIL is a useful language learning method. This has been shown by as yet scanty but gradually increasing research on CLIL in Finland. Laitinen (2001) found that her grade 5 immersion subjects (11–12 years of age) did far better than the average grade nine students (15–16 years of age) in the national comprehensive school test of English. Other small-scale studies show similar results.

#### *Pedagogical implications*

Attainment of near-native comprehension skills seems to have resulted in stagnation or plateau of language development in French Immersion in Canada. In French Immersion, attempts to trigger stagnated language development have been suggested, such as administering language learners *challenging production tasks* (output hypothesis, Swain 1993) and ‘*negotiation of form*’ tasks. It seems that formal language instruction should be practiced and less time could be spent on comprehension exercises, especially at elementary levels. Instead, accuracy practice (‘negotiation of form’ tasks) could be devoted more time in formal language classes, as well as *authentic*, up-to-date language, such as newspapers, *literature*, registers, pragmatic factors, and cultural aspects. For optimal results, language and content teachers would need to *collaborate* in planning, implementing and evaluating integrated units. In extensive implementation of CLIL, collaboration at the level of syllabi would undoubtedly lead to profound restructuring and re-organization.

In order to enhance language production, CLIL should contain more spoken production, not only by individual students but also by pairs and groups of students. Teachers should focus on *instructional language* demanded by e.g. *explanations* and ‘*quality*’ *questions* which require use of *higher-order thinking skills* (Mohan 1986) and in-depth-processing, and elicit longer responses and more language from students. The amount of *student interaction* in class could be increased. Interactive tasks could be planned with focus on CLIL, in order to enhance language development, strategy development, and content-specific or general thinking skills and related knowledge structures.

## References

- Berent, Gerald P.  
1994 The subset principle in second-language acquisition. In *Research Methodology in Second-Language Acquisition*, Elaine E. Tarone, Susan M. Gass and Andrew D. Cohen (eds.), 17–39. Hillsdale, NJ: Lawrence Erlbaum.
- Bley-Vroman, Robert and Craig Chaudron  
1994 Elicited imitation as a measure of second-language competence. In *Research Methodology in Second-Language Acquisition*, Elaine Tarone, Susan Gass and Andrew Cohen (eds.), 245–262. Hillsdale, NJ: Erlbaum.
- Brinton, Donna M., Marguerite A. Snow and Marjorie B. Wesche  
1989 *Content-Based Second Language Instruction*. New York: Newbury House.
- Chomsky, Noam  
1996 *The Minimalist Program*. Cambridge, Massachusetts: MIT Press.
- Common European Framework of Reference for Languages:  
2001 *Learning, teaching, assessment*. Council of Europe: Cambridge University Press.
- Cummins, Jim  
1984 *Bilingualism and Special education: Issues in Assessment and Pedagogy*. Clevedon, England: Multicultural Matters.
- Cummins, Jim and Merrill Swain  
1986 *Bilingualism in Education*. London: Longman.
- DeKeyser, Robert M.  
2001 Automaticity and automatization. In *Cognition and Second Language Instruction*, Peter Robinson (ed.): 125–151. Cambridge: Cambridge University Press.
- Doughty, Catherine  
2001 Cognitive underpinnings of focus on form. In *Cognition and Second Language Instruction*, Peter Robinson (ed.), 206–257. Cambridge: Cambridge University Press.
- Doughty, Catherine and Jessica Williams  
1998a Issues and terminology. In *Focus on Form in Classroom Second Language Acquisition*, Catherine Doughty and Jessica Williams (eds.), 1–11. Cambridge: Cambridge University Press.  
1998b Pedagogical choices in focus on form. In *Focus on Form in Classroom Second Language Acquisition*, Catherine Doughty and Jessica Williams (eds.), 197–261. Cambridge: Cambridge University Press.
- Ellis, Nick  
1994 Introduction. Implicit and explicit language learning-an overview. In *Implicit And Explicit Learning Of Languages*, Nick Ellis (ed.), 1–31. London: Academic Press.

- Ellis, Nick  
1995 Consciousness in second language acquisition: a review of field studies and laboratory experiments. *Language Awareness* 4 (3): 123–146.  
2001 Memory for language. In *Cognition and Second Language Instruction*, Peter Robinson (ed.), 33–68. Cambridge: Cambridge University Press.
- Flynn, Suzanne  
1996 A parameter-setting approach to second language acquisition. In *Handbook Of Second Language Acquisition*, William C. Ritchie and T. K. Bhatia (eds.), 121–158. San Diego, California: Academic Press.
- Flynn, Suzanne and I. Espinal  
1985 Head-initial/head-final parameter in adult Chinese L2 acquisition of English. *Second Language Research* 1 (2): 93–117.
- Fruhauf, Gianna, Doy Coyle and Ingeborg Christ (eds.)  
1996 *Teaching Content in a Foreign Language: Practice and Perspectives in European Bilingual Education*. Alkmaar: The European Platform for Dutch Education.
- Gass, Susan  
1988 Integrating research areas: a framework for second language studies. *Applied Linguistics* 9 (2): 198–217.  
1994 The reliability of second-language grammaticality judgments. In *Research Methodology in Second-Language Acquisition*, Elaine Tarone, Susan Gass and Andrew Cohen (eds.), 303–322. Hillsdale, NJ: Erlbaum.
- Genesee, Fred  
1987 *Learning Through Two Languages*. Cambridge MA: Newbury House.
- Grabe, William and Fredricka L. Stoller  
1997 Content-based instruction: Research foundations. In *Content-Based Classroom. Perspectives on Integrating Language and Content*, Ann M. Snow and Donna M. Brinton (eds.), 5–21. White Plains, N.Y.: Addison-Wesley Longman.
- Gregg, Kevin R.  
2001 Learnability and second language acquisition theory. In *Cognition and Second Language Instruction*, Peter Robinson (ed.), 152–180. Cambridge: Cambridge University Press.
- Hawkins, Roger  
1989 Do second language learners acquire restrictive relative clauses on the basis of relational or configurational information? The acquisition of French subject, direct object and genitive restrictive relative clauses by second language learners. *Second Language Research* 5 (2): 156–188.
- Krashen, Stephen  
1985 *The Input Hypothesis: Issues and Implications*. London: Longman.
- Kristiansen, Irene  
1990 Nonverbal Intelligence and Foreign Language Learning. *Research Bulletin* 73. Department of Education. University of Helsinki.



- Lapkin, Sharon, Merrill Swain and Susan Shapson  
1990 French immersion research agenda for the 90s. *Canadian Modern Language Review* 46, 638–674.
- Laitinen, Jaana  
2001 English immersion in Finland – dreams or reality ? A case study on the Hollihaka early total English immersion programme, its structure, implementation and learning outcomes. Licentiate diss., University of Jyväskylä. Centre for Applied Language Study.
- Long, Michael H. and Peter Robinson  
1998 Focus on form: theory, research, and practice. In *Focus on Form in Classroom Second Language acquisition*, Catherine Doughty and Jessica Williams (eds), 15–41. Cambridge: Cambridge University Press.
- Lust, Barbara, Suzanne Flynn and Clare Foley  
1996 What children know about what they say: elicited imitation as a research method for assessing children’s syntax. In *Methods For Assessing Children’s Syntax*, 55–76. Cambridge, MA: The MIT Press.
- Marsh, David and Bruce Marsland (eds.)  
1999 CLIL Initiatives for the Millennium. Report on The CEILINK Think-Tank. University of Jyväskylä. Continuing Education Centre.
- Mohan, Bernard A.  
1986 *Language and Content*. Reading, MA: Addison-Wesley.
- Myles, Florence  
1995 Interaction between linguistic theory and language processing. *Second Language Research* 11 (3): 235–266.
- Nikula, Tarja and David Marsh  
1996 Kartoitus vieraskielisen opetuksen tarjonnasta peruskouluissa ja lukioissa. Opetushallitus. [A survey of mainstream bilingual education in the Finnish general education. National Board of Education.]  
1997 Vieraskielisen opetuksen tavoitteet ja toteuttaminen. Opetushallitus. [Content and language integrated learning in the Finnish primary and secondary sectors: from aims to implementation. National Board of Education.]
- Sawyer, Mark and Leila Ranta  
2001 Aptitude, individual differences, and instructional design. In *Cognition and Second Language Instruction*, Peter Robinson (ed.), 319–353. Cambridge: Cambridge University Press.
- Schmidt, Richard  
1990 The role of consciousness in second language learning. *Applied Linguistics* 11 (2): 129–158.  
2001 Attention. In *Cognition and Second Language Instruction*, Peter Robinson (ed.), 3–32. Cambridge: Cambridge University Press.

- Snow, Marguerite A. and Donna M. Brinton (eds.)  
1997 *The Content-Based Classroom. Perspectives on Integrating Language and Content.* NY: Addison Wesley Longman.
- Stern, H. H.  
1985 Bilingual schooling and foreign language education. In *Perspectives on Bilingualism and Bilingual Education*, J. E. Alatis and J. J. Staczek (eds.), 399–422. USA: Georgetown University Press.
- Swain, Merrill  
1985 Communicative competence: some roles of comprehensive input and comprehensible output in its development. In *Input in Second Language Acquisition*, Susan Gass and C. Madden (eds.), 235–253. Rowley, MA: Newbury House.  
1993 The output hypothesis: just speaking and writing aren't enough. *The Canadian Modern Language Review* 52: 158–164.
- Swain, Merrill and Sharon Lapkin  
1982 *Evaluating Bilingual Education: A Canadian Case Study.* Clevedon: Multilingual Matters.
- Takala, Sauli  
1997 Introduction. In *Mainstream Bilingual Education in the Finnish Vocational Sector*, David Marsh, Paula Oksman-Rinkinen and Sauli Takala (eds.), 9–21. Helsinki: Opetushallitus [Board of Education].
- Tomlin, Russell S. and V. Villa  
1994 Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition* 16 (2): 183–203.
- Wesche, Marjorie B.  
2001 Editorial. *The Canadian Modern Language Review. French Immersion and Content-based Teaching in Canada. Special Issue* 58 (1): 1–8.  
2002 Early French immersion: how has the original Canadian model stood the test of time? In *An Integrated View of Language Development*, Petra Burmeister, T. Piske and A. Rohde (eds.), 357–379. Trier: Wissenschaftlicher Verlag Trier.
- Wesche, Marjorie and Peter Skehan  
2002 Communicative, task-based, and content-based language instruction. In *The Oxford Handbook of Applied Linguistics*, Robert Kaplan (ed.). New York: Oxford University Press.
- White Paper on Education and Training  
1995 *Teaching and Learning – Towards the Learning Society.* Brussels: The European Commission.
- Winter, Bill and Arthur S. Reber  
1994 Implicit learning and the acquisition of natural languages. In *Implicit and Explicit Learning of Languages*, Nick Ellis (ed.), 115–145. London: Academic Press.

# **Effects of teacher discourse on learner discourse in a second language classroom**

*Tsuyoshi Kida*<sup>1</sup>

## **1. Introduction**

This chapter will focus on the issue of the acquisition<sup>2</sup> of spoken discourse competence by second language learners in an instructional environment. Attention will be paid more specifically to how the teacher's discourse motivates the learner's discourse in a classroom, and how the language teacher contributes to the development of the learner's discourse competence, that is, the ability to organize one or more utterances in an appropriate fashion (cohesiveness and coherence) in a given situation, and knowledge of discourse structures and markers. These skills are especially important for conducting argumentative discourse with long or complex utterances, or for producing narrative discourse. Indeed, learners need to improve this competence, without which they would have practical, social, or even psychological problems in their host community (e.g. integration in the community). This study will be limited to the context of classrooms where French is being taught as a second language to adult learners who have come from various countries to France. The objective of the study is to underline some of the problems encountered in such a classroom by describing what occurs in class and how classroom discourse is dealt with by the participants. The question raised is: what conditions might allow learners to benefit from classroom interactions so that it will improve their discourse competence.

## **2. Framework of the study**

The classroom observed for this study mainly involved naturally occurring communicative interaction (conversation). This pedagogical approach is designed to remedy the weaknesses of the so-called communicative method, which fails to deal with learners' performance in unpredictable situations, and to offer them genuine experience with the communicative and

social practices of language (see Moirand 1990), viewed as crucial for the acquisition of communicative competence (Widdowson 1981; Hymes 1984). The aim of this chapter is to find out whether discourse acquisition is possible in such a conversation-based classroom by illustrating the discourse and interactional activities of the participants.

The issue of discourse acquisition is probably more difficult to approach than that of grammar (or linguistic) acquisition, given that discourse has fewer explicit rules than the grammar. Whatever the case may be, little is known about the discourse acquisition process, even less than about grammar acquisition. Swain's (1985) study on the Canadian French immersion program suggests a difference in acquisitional outcomes between discourse and grammar. Her survey revealed that after seven years of French immersion, the students had attained a native-like level of (written) discourse competence in spite of their problematic usage of morphology and syntax. That is, the acquisition of grammar might be different from that of discourse (Swain 1985: 242 and 252). Interestingly, a similar argument has been advanced in neurolinguistics claiming that the context-sensitive use of discourse and the non-contextual use of linguistic rules are different in nature (Paradis 1998)<sup>3</sup>. It could be that the ability to produce discourse in social events is more complex than acquiring the procedural and declarative knowledge on which linguistic competence is based (Cohen 1984)<sup>4</sup>.

If we assume that discourse is a specific object that second language learners must master, then several issues must be dealt with before we can start investigating discourse learning in the classroom. One of them might be the need to determine what discourse competence learners have as native speakers of their mother tongue or other language(s), and how this competence is used or transferred when they produce discourse in a second language. This issue would call for a careful contrastive analysis of a learner's discourse performance and a longitudinal collection of data. This is beyond the scope of this chapter which is aimed at observing a second language classroom.

Another issue we could explore is how the learners in a classroom deal with discourse. That is, are there specific materials or other sources of input offered by the teacher to the learners for their discourse learning in the classroom? Are the learners able to test their hypotheses on discourse knowledge or to process it in a satisfactory fashion for their discourse acquisition in the classroom? These questions can be reformulated in the terms that have been central in second language acquisition research, namely, in terms of input and output. Presumably, this way of framing the problem helps us to conceive of ways of observing classrooms, and it is

necessary to start with the question of how – in terms of input or output – discourse learning takes place in an instructional environment.

Considering the input issue presupposes raising some theoretical questions about discourse as input for discourse learning in the classroom; this ultimately brings us to the question of what discourse is. It also requires elaborating a pedagogical norm stipulating what discourse rules should be learned, such as markers, structures, cohesion or coherence rules, some rhetorical conventions proper to the target language (e.g. how to use a repetition), and to reflect upon the way input can be offered to learners. Nevertheless, the learning context analysed in this study, namely a French second language class in France, is not designed to offer such a program to learners. Empirical observation of several hours of conversation-based classroom data did not provide any clear evidence that the teachers' input pertained to the learners' discourse knowledge or competence. It was mostly limited to linguistic problems in the learners' discourse by means of negotiation of meaning, corrective feedback, modified interaction, or metalinguistic comments. The teacher's discourse in classes like these tends to display typical "foreigner talk" features such as overly elicitive questions to learners (Long and Sato 1983; Lörcher 1986), over-scaffolding repetition (Faraco 1995, 2002)<sup>5</sup>, or elliptical discourse with nonverbal emphasis (Kida forthcoming). Thus, the types of discourse input to which learners are exposed in the classroom are limited to simplified discourse, at least in the instructional context analyzed in this study. Therefore, the issue of input will not be further addressed in this study. For a study of discourse learning by means of classroom input, see House (1996).

If learners cannot expect efficient discourse input from classroom interactions, it can be assumed that they learn discourse outside the classroom. However, learners' "careers" in discourse acquisition are still not sufficiently documented to take this assumption for granted. It would be of interest to comparatively investigate the way in which learners gain discourse experience inside and outside the classroom. The present study could not explore such an ethnolinguistic perspective because of the nature of the data available here. But, according to the testimonies of our learners, their discourse "lives" seem to be very limited. Some said they did not have any native-speaker friends with whom they could converse on a daily basis; others had such friends or non-native classmates, but their conversations were usually of a spontaneous nature or were in their mother tongue. Most of them, unlike immigrant workers, are rarely exposed to the environment of institutionalised discourse and have only the mass media or their own reading as sources of extended discourse. Therefore, it is assumed in

this study that the learners had minimal experience with discourse input outside the classroom.

The issue of output (Swain 1995) allows for different perspectives than the issue of input. Here it is necessary to analyse the conditions under which the learners produce an extended or complex discourse. For this issue, it is legitimate to assume that discourse acquisition requires experiencing communicative and social execution of discourse, as Hymes (1984) suggested for the acquisition of communicative competence. Learners' discourse output in a conversational classroom does not seem to be very elaborate, as Swain (2000: 99) pointed out: "[...] although students used French in class, little of it included extended discourse [...]." However, are students always reluctant to produce extended discourse? This is one of the questions tackled in this study.

In fact, in any interaction, the dynamics can change in such a way that the utterances used become more complex and participation in the interaction becomes more active. If the teacher's discourse and classroom interaction could motivate the learners to use a complex discourse rather than the simple productions employed in language acquisition sequences (see Faraco & Kida 1999), then they could engage in producing at least some output that would promote their discourse acquisition. In this perspective, the present chapter will attempt to highlight the conditions conducive to discourse motivation in the classroom, and to explore one of the ways of analysing teacher/learner interactions favouring discourse learning in an instructional context.

### **3. Research questions**

As outlined earlier, the questions to be considered in this study are the following:

- 1) Do classroom interactions influence the learners' discourse? If so, what types of teacher discourse are favourable to the learners' experience of more elaborate discourse?
- 2) Is classroom interaction at all useful for discourse learning? In other words, is discourse acquisition possible in a conversation-oriented classroom?

To study the effects of teacher discourse on learner discourse, learner discourse will be examined in relation to the interactional context and the pragmatic orientation of teacher discourse. Two types of data analysis are

proposed: (1) a quantitative analysis of occurrences of some features of the learners' and teacher's discourse, in order to draw a general picture of teacher/learner interactions in the classroom (i.e. a discourse analysis); (2) a qualitative analysis of interaction contexts in the classroom that might not show up in the discourse analysis (i.e. a context analysis). Given that the context analysis is derived from traditional conversational analysis, a detailed methodological explanation is not required here. For the discourse analysis, the next section will describe the specific methodology adopted in this study.

#### **4. Data and methodology of the study**

##### **4.1. Database**

Data samples are transcribed from a one-hour videotaped recording of a conversation-oriented class in French as a second language that took place in Aix-en-Provence (France) at the end of the semester (June 2000). The recording is a pilot recording for a broader classroom research project aimed at analysing the nonverbal behaviour of classroom participants, which will be reported in subsequent papers (Faraco and Kida forthcoming). The participants were the teacher and seven adult learners (one Catalan, one Chilean, one Japanese, two Koreans, one Vietnamese, and one Cypriot with Greek as the mother tongue). The learners were all close to thirty years of age (the Koreans and the Cypriot had been twenty-year students in their respective countries). They were all attending the same intermediate-level class but, as is the case in any language class, their proficiency levels varied. Note that normally their class was comprised of fifteen learners, but the number of participants had to be limited to eight for technical reasons, although most learners expressed the desire to attend to the experimental lesson. The class was thus divided into two groups, one of which is analysed in this study. The recording did not take place in their regular classroom. However, the teaching style and the relations among the participants seemed to be the same as during their regular lessons, and they felt that it was a lesson like any other despite the uncommonness of the situation. Only the teacher stated that she was more concerned with grammatical correctness than in regular lessons, since the learners were preparing for an exam to be held a couple of days later.

A one-hour recording of a classroom is clearly not enough to fully analyse activities relating to the learning and teaching of discourse, for



there are always many different factors and variations of discourse and interaction within a given classroom. The aim of this study was nonetheless to bring out the interactional dynamics of the class and within these dynamics, to highlight some principles of communicative motivation conducive to complex discourse. In this sense, it was posited that a set of data samples from a single lesson given by one teacher is worthy of analysis.

#### 4.2. Pragmatic and interactional context

The first objective is to describe the discourse context of learners' utterance in terms of turn attribution and pragmatic constraints. In fact, learners' utterances are produced in different contexts of interaction: firstly, whether an utterance follows the teacher's turn or that of another learner; secondly, whether the teacher's discourse is addressed either to a specified learner (named or not) or to all participants; and thirdly, what type of pragmatic orientation the previous utterance implicates. Moreover, some learners' utterances do not enter into any of these cases, especially when they are produced out of turn attribution or when several learners simultaneously try to take the floor. A speaking turn can include several utterances. The pragmatic and interactional context of each segment of learner discourse is analysed by considering all of these cases.

Interactional contexts were labelled as "Alter", "Collective", "Individual", "Learner", and "Own utterance". "Alter" means that an utterance was produced in a turn attributed to a learner other than the speaker of that utterance. "Collective" was used to label a teacher's utterance that did not specify any next speaker. "Individual" was used when the current speaker was explicitly or implicitly indicated by the preceding utterance. In general, the use of certain French pronouns is the best way to distinguish "Individual" (*tu*) from "Collective" (*vous* or a third-person pronoun referring to a previous speaker). Other distinctive features determining the Collective context were the co-occurrence of a confirmation question (e.g. *d'accord?*), or a long utterance with explicative value. The gaze and posture of the speaker can also be taken into account to differentiate these types of turn attribution. "Learner" indicates that the utterance followed the turn of another learner, and designates exchanges between two learners. Finally, because a learner's turn can be made up of several utterances, "Own utterance" refers to the second or a later utterance in the turn.

To establish a typology of teacher discourse, the categories adopted were as follows: 1) Message seeking, 2) Message checking, 3) Information



question, 4) Modified interaction, 5) Feedback, 6) Explication, and 7) Assertion. This typology is based on Schiffrin's (1994) discourse analysis system, which describes the main types of elicitation in interview settings (see also Moder and Halleck 1998: 119–121), and then completed as the data were analysed. The typology was also intended to be general enough to be used in analyses of other types of situations, such as interactions outside the classroom.

Message seeking involves various forms of questions (sometimes indirect) whereby the speaker attempts to get unknown information from a participant (e.g. *Qu'est-ce que c'est ce dont vous êtes fier?* 'What are you proud of?'). Message checking differs from Message seeking in that the latter is about an unknown referent, whilst the former is used either to verify information given on a previous turn, especially when the previous utterance was totally or partly ambiguous (e.g. *Qu'est-ce que ça veut dire?* 'What does that mean?'), or to confirm whether the current speaker had adequately identified a reference or an intention in the interlocutor's message (e.g. *Laquelle télévision est culturelle, la coréenne ou la française?* 'Which TV is cultural, Korean or French?'). Information question consists of a simple question (e.g. question about grammatical knowledge) as frequently found in classroom settings (e.g. *Quel est le substantif de fier?* 'What is the noun form of fier?'). Another device used is corrective feedback by reformulation or Modified interaction (see Long 1996), which differs from Message checking in that the former focuses on a linguistic topic, whilst the latter is about a conversational topic. As we shall see, this difference may have a crucial effect on learner discourse. For Feedback and Explication, often attested in teacher discourse, the former pertains to when the teacher maintains an extended exchange with a learner, and the latter occurs when the teacher explains linguistic or cultural information (e.g. *Fier, c'est en principe faire quelque chose...* 'Fier, it's usually doing something...'). Finally, other utterances with no elicitation value were coded as Assertion.

#### 4.3. Features of learner discourse

One of the methodological difficulties encountered was how to quantitatively measure learners' discourse. It was assumed that when a discourse is long, highly marked, topicful, and subjectively involved, the learners are more inclined to produce discourse and more are motivated by the teacher's discourse. Hence, quantifiable items were sought first, with careful attention to the context in which each utterance was made. Such a contextual analysis

(see Kendon 1990) is necessary, in particular to determine whether the ongoing topic is developed with new information, and whether the learners' turn-taking behaviour is active in class participation. Note that strategic cues (e.g. pause fillers) were not taken into consideration in this study since they could not always be viewed as discourse strategies. The analysis of other qualitative aspects (coherence, cohesiveness, structural complexity, discourse organization) will be reported in the context analysis (Section 6).

The following features made up the first series of criteria: 1) mean length of utterance (MLU), 2) number of discourse connectors, 3) number of utterance modulators, 4) Turn initiation, and 5) Topic advancement. In order to properly calculate MLU, some specific phenomena in French (e.g. elision, word contraction) must be briefly considered first. The utterance *J'ai compris Pierre* ('I understood Pierre'), which seemingly only contains three "words" because of the elision of the "e" in "Je", was counted as a four-word utterance by analogy of *Il a compris Pierre* ('He understood Pierre'). By contrast, *parce que* ('because') was considered a single word like *puisque* ('since'), and *tout le monde* ('everybody') was counted as one word.

Regarding discourse connectors, not only are classic conjunctions counted but also pre-utterance pronouns like *moi, ...* ('me,...') or *là, ...* ('there,...'), the non-responsive *oui*, pre-utterance adverbs like *maintenant* ('now') and other expressions that act as introducers like *bon* ('well'), *en fait* ('indeed'), or *par exemple* ('for example'). Utterance modulators are expressions of the speaker's subjective involvement vis-à-vis his or her utterance or traces of "analytic interpretation of fact" (Wertsch 1991). They involve expressions like modal devices (e.g. *je pense que...* 'I think...', *je crois que...* 'I believe...'), adverbs of personal opinion or similar adverbs (e.g. *personnellement* 'personally', *d'après moi* 'according to me', *pour moi* 'for me'), expressions of disengagement (e.g. final particle *quoi* 'what', final non-locative particle *là* 'there'), strong engagement, affiliation to the conversational context (e.g. *moi aussi* 'me too'), factual approximation or imprecision (e.g. *comme ça* 'like that', *quelque chose de...* 'something...', ... *tout ça* '...all that', mid-utterance *par exemple* 'for example'), dramatic *mise en scene* like direct reported discourse, and a number of emphatic devices like repetition and nonverbal modulation (intonation and gesture). It should be noted with regard to repetitive modulation that repetition due to a systemic obstacle – as caused, for example, by overlapping speech – and hesitant repetition were not considered as modulators; only repetitions seemingly comprising a certain discursive function like emphasis were seen as modulators. As regards direct reported discourse, which itself is

considered to be a kind of modulation, various devices in this discourse (intonative modulation, connectors, etc.) are not counted as such.

The second set of criteria for determining the types of learner discourse were the following: 1) Feedback, 2) Question, 3) Response, 4) Reply, and 5) Modified output, following the teacher's (in some cases, a learner's) corrective utterance. These criteria are simpler than those used to describe the teacher's discourse so as to allow for an open interpretation. Discourse types which did not enter into these categories were assessed only in terms of the first set of items described above. These discourse types were used as complementary criteria related to turn and topic (for example, a reply and a question are more topicful than a feedback; a response and a modified output are produced in an other-initiated turn). They are also related to discourse length. For instance, feedback or modified output is shorter than questions, responses, or replies. As seen below, these discourse types and features are more or less context-dependent.

In principle, there is no link between discourse length or complexity and its qualitative aspects (for example, good argumentation). Nevertheless, learners are reluctant to produce discourse in the classroom and the working hypothesis is that a quantitative increase in discourse complexity in their utterances could be the first sign of a class that is rich in discourse learning. Furthermore, the context analysis was expected, at least partly, to compensate for what the discourse analysis could not fully demonstrate. These two analyses were intended to provide different pictures of the same data samples, and to capture classroom interactions in a complementary manner. On the other hand, the discourse analysis might have the advantage of making it possible to formulate some practical suggestions for improving teacher discourse. Whether a combination of context and discourse analyses could reveal something about the dynamics of classroom interaction could be a methodological issue in conducting classroom interaction research.

## **5. Study I: Discourse analysis**

### **5.1. Interactional context variation**

Table 1 shows the general tendency of all learners' discourse in terms of mean length of utterance (MLU), Connector and Modulator, Turn initiation, Topic advancement, and discursive types such as Feedback, Question, Response, Reply, and Modified output.

*Table 1.* General distribution of learners' discourse features and types

	Utter.	MLU	Connec.	Modul.	Turn initiat.	Topic advan.	Feedback	Question	Response	Reply	Modif. output
Occur.	512	2909	229	313	156	250	96	30	139	26	31
per utter.	1.00	5.68	0.45	0.61	0.30	0.49	0.19	0.06	0.27	0.05	0.06

The first row gives the total number of occurrences of each item, and the second row gives the mean number of occurrences per utterance. The columns MLU, Connector and Modulator were calculated cumulatively in terms of words, and the other columns were using the binary system (0 or 1), in terms of occurrences. Some of the items led to a theoretical quantification, and although they do not refer to a discursive reality, they are useful in comparisons.

The table indicates that the mean length of utterance was 5.68 words. Note that the real length could be slightly longer than this score, since some nonverbal turns were included in the number of utterances. Connectors and modulators appeared at a rate of 0.45 and 0.61 occurrences per utterance respectively, which means approximately one connector or modulator every two utterances in the learner's discourse. As for the types of learner discourse, the learners initiated 30% of the utterances without turn attribution, and 49% of the utterances included a topic advancement, that is, additional information. An utterance constituted a feedback at a rate of 19%, a question at 6%, a response at 27%, a reply at 5%. Modified output was attested in 6% of all learners' utterances.

This configuration of learner discourse varied with the turn-attribution or interaction context (see Figure 1). As for the frequency of the different interactional contexts, the main contexts, in decreasing order of frequency, were as follows: Individual (161), Own utterance (132), Collective (106), Learner (87), and Alter (13). Note that the frequency order of the types of interaction followed the order of clarity of turn assignment. So the principle is as follows: the more clearly a turn is attributed, the more easily the learners participate in exchanges. The rareness of the Alter context can be explained by the fact that the speaker in this context must go against the ongoing course of the exchange by deliberately taking a turn or by strongly initiating a topic. The learners probably hesitated about such a behaviour,

for they run the risk of appearing unsociable vis-à-vis other participants, particularly in a situation like a second language classroom. Maybe this tendency would be different outside the classroom.

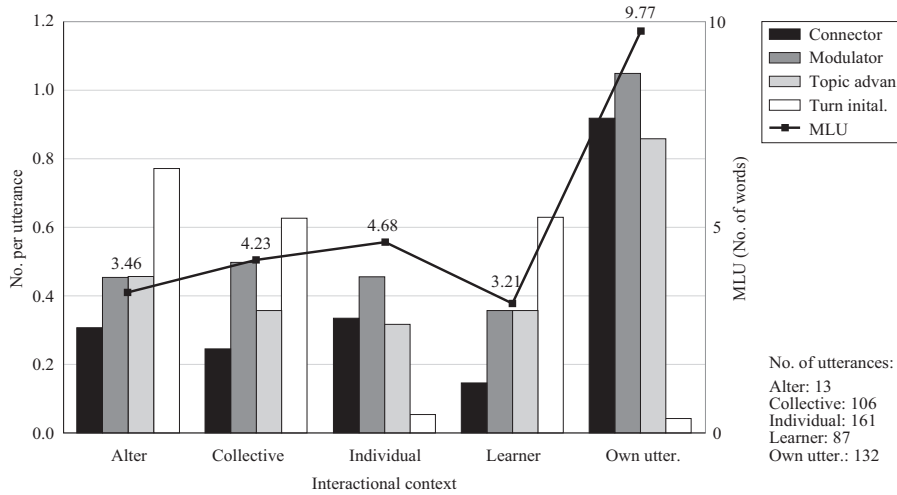
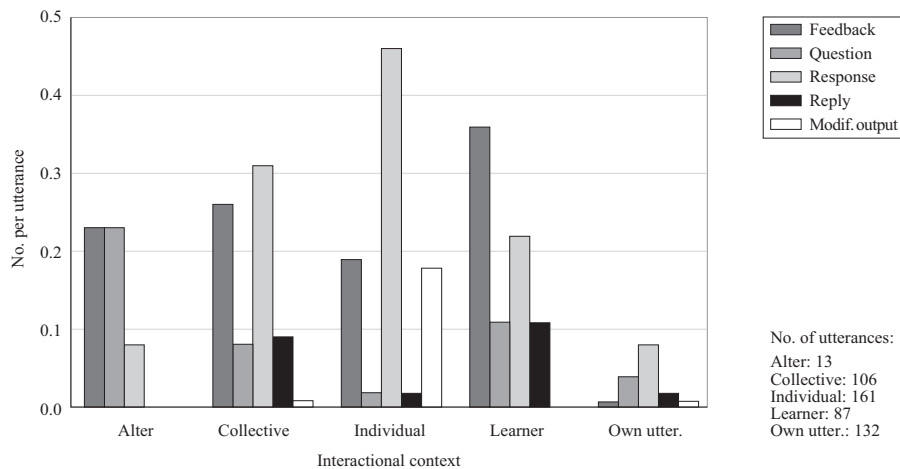


Figure 1. Variations in learner discourse features across interactional contexts

As for the quantitative aspect of learner discourse, Figure 1 reveals that utterances in the Own utterance context are more than twice as long, marked and topicful than in the other contexts, which were comparable to each other. This could mean that the learners tended to produce more complex discourse after the first utterance in a given turn, and that their first utterance was equally constrained by the prior utterance of the last speaker, no matter how the turn was assigned. This quantitative tendency in the learner discourse was not the same at the qualitative level, however. When another learner was the previous speaker (Learner context), the learners seemed to use fewer connectors than in the Alter, Collective, and Individual contexts. That is, an exchange between learners was less likely to promote the use of discourse markers. Finally, it turned out that turn initiation was more context-sensitive than the other criteria. It was found less often in the contexts where other criteria had higher scores (Individual and Own utterance contexts). This point will be taken up again in the discussion.

Figure 2 represents the variation of the types of learner discourse across interactional contexts. Although subtly, it seems to indicate that all types of discourse were more or less context-dependent. Modified output occurred

exclusively in the Individual context. Feedback was the most frequent discourse type in the Learner context; this tendency is consistent with the fact that the learner discourse (at least the first utterance) tended to be less elaborate after a learner had taken the floor (see Figure 1). Questions were asked less frequently, especially when the turn was in the Individual or Own Utterance context, which were, respectively, highly other-initiative, or self-initiative for another purpose such as a topic pursuit. It might be that for a learner to easily ask a question, the context must be more or less free from any constraints on the next utterance, like the Alter, Learner, or Collective contexts.



*Figure 2.* Variations in learner discourse types across interactional contexts

Response was also a context-dependent discourse. It was widely given in the Individual context, but also to a certain degree in the Collective and Learner contexts. This discourse type indeed needs a rationale and it is certain that the context where the next speaker was specified (Individual context) provided such an occasion better than other contexts. Figure 2 indicates that this occasion was still possible in the Collective and Learner contexts but less probable than in the Individual context. Response in the Alter context means that another learner responds instead of the learner to whom the teacher addressed a question. Response in the Own utterance context corresponds to the case where a learner answers in the second utterance within a speaking turn after stating a preliminary in the first utterance. Those cases were nevertheless rare, as Figure 2 suggests. Finally, the low frequency of all discourse types in the Own utterance context means that in this context,

learners used other types of discourse than the ones included here, such as, for example, assertive discourse or other types. Thus, the data suggests that the type of interaction context and the type of learner discourse are generally interdependent.

This overall analysis provided a rough characterization of learner discourse based on the interactional context. One can obtain a finer view of the influence of context by examining the effect of the teacher's discourse and that of the learner's discourse on the learner's discursive activity. This part of the study will focus on characterizing learner discourse in terms of the teacher's talk, since there is room for teacher intervention. Accordingly, the next section will attempt to give a more detailed description of the effect of the teacher's discourse in individualized exchanges with learners, that is, the relationship between the learner's discourse and the pragmatic orientation that the teacher's discourse implicates in the Individual context.

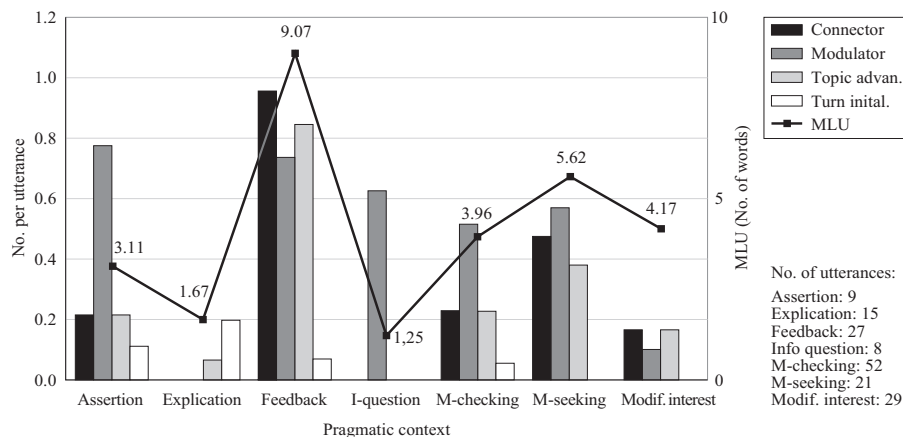


Figure 3. Variations in learner discourse features across pragmatic contexts (Individual context)

## 5.2. Teacher's discourse in the Individual context

The different types of teacher discourse in the Individual context are shown as a function of the discursive features of the learners' utterances in Figure 3, and of the learner discourse types in Figure 4. The features and type of data are commented upon together here since they are correlated, as seen in the previous subsection. As these figures show, the type of the

teacher's discourse differentiated between the learners' discourse performance. In Figure 3, note that the context that created the longest utterance length (MLU), and the most markers (connectors and modulators) and topic advancements in the learner discourse was the teacher's feedback. This finding seems to be in line with the finding in the last section concerning the fact that the Own utterance context was the most productive generator of discourse complexity (see Figure 1), insofar as teacher Feedback may not be perceived as an autonomous turn. That is, the learners probably felt that their discourse was as continuous as in the Own utterance context. By contrast, the Information Question context, to which only a short response is expected, and the Explication context, where there was no particular elicitation value, seem to be less discourse-rich, at least in terms of MLU. These findings are partially correlated to those in Figure 4, which shows that the Explication context was strongly related to the learners' feedback, which was a less complex kind of discourse.

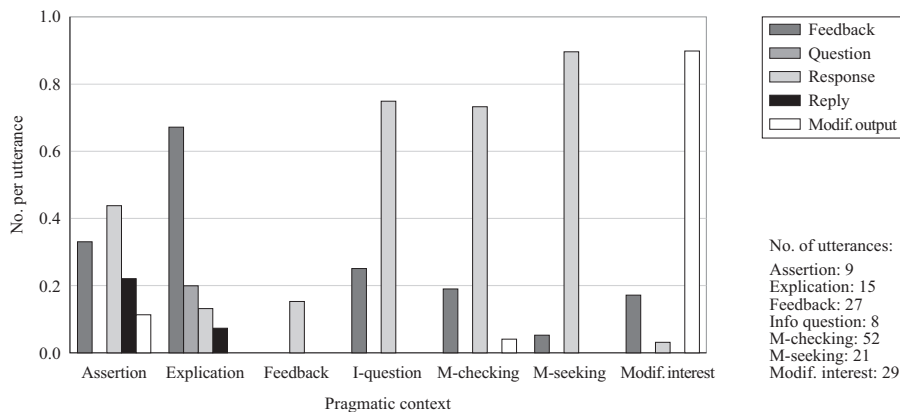


Figure 4. Variations in learner discourse types across pragmatic contexts (Individual context)

The other contexts are not easy to interpret. Figure 4 indicates that among them, Message seeking scored higher than Assertion, Message checking and Modified interaction, which were relatively comparable in terms of MLU. Modified interaction is easier to analyse. Figure 4, as expected, indicates a strong link between Modified interaction and Modified output, which does not require discursive cues or subjective involvement on the part of the addressee. Therefore, it seems that Modified interaction has an underlying principle that differs from that of the other discourse contexts.



Figure 3 indicates that the effect of the teacher’s discourse type on the learners’ utterances was relatively stable in terms of modulators (except for Modified interaction). In contrast, the use of connectors was linked more to Message seeking than Message checking or Assertion. That is, the role of the teacher’s discourse in the use of discourse markers by learners is greater for connectors than for modulators. Interestingly, this is similar to the tendency observed in the turn-attribution typology (see Figure 1). As for topic advancement, this was influenced by the teacher discourse type in the same way as were discourse connectors.

To sum up, compared with other teacher discourse types, Message seeking, after Feedback, seems to be a context that is most favourable to the use of discourse markers, longer lengths and many topics in learners’ discourse. Admittedly, if speakers are completely free from discourse constraints such as a given context after another’s assertion, they must contextualize their discourse with a short introductory utterance. On the other hand, the utterance immediately following after a message checking might be short, as in a backchannel, a repetition, or the like. The learners, then, might benefit from an average regulation of context to structure a longer, and more highly marked and topicful discourse.

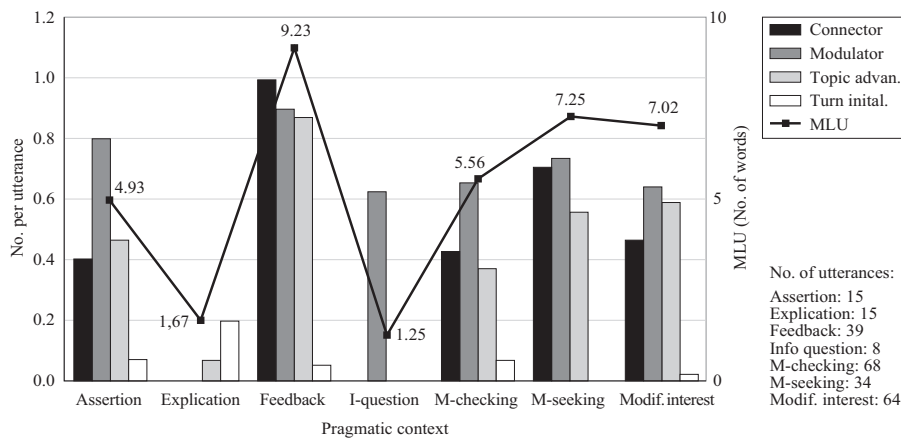


Figure 5. Variations in learner discourse features across pragmatic contexts (Individual context combined with Own utterance context)

### 5.3. Multi-utterance tendency

The analysis so far has focused only on the first utterance of a learner's turn. However, a turn sometimes includes several utterances. The above data sampling excluded subsequent utterances, which were counted in the Own utterance context. This methodological shortcoming needs to be overcome. Figure 5 attempts to recapitulate all the utterances within the learners' turns that occurred next to the teacher's discourse. A discourse-type analysis is not conducted here, since the effect of teacher discourse on learner discourse types after the first utterance would make little sense.

Compared with Figure 3, Figure 5 offers a slightly different picture. Higher values in Figure 5 than in Figure 3 indicate that certain parts were affected by including utterances in the analysis. This effect, as argued above, did not occur in modulation, except in the Modified interaction context. The Assertion, Message checking, Message seeking and Modified interaction contexts showed an increase in MLU, connectors, and topic advancement, which means that these discourse contexts had a more constraining effect on the first learner utterance than on the later utterances with regard to utterance length, discourse markers and topics. The fact that the Explication, Feedback, and Information question contexts did not exhibit a significant effect on any feature indicates that the tendency in Figure 3 still holds true within speaking turns. Despite the added data, the general tendency is similar. Thus, Feedback is still the context that created the richest discourse complexity. Explication and Information Question created less discourse, and Message Seeking played a greater role in the learners' discourse performance than the Message Checking, or Assertion contexts.

However, the Modified interaction context exhibited a striking increase on all criteria examined here, and, especially in the mean length of utterance. In other words, although this context powerfully constrains the learner's next utterance, as Figure 4 has indicated, it does give rise to additional utterances after the first one. That is, Modified interaction does not discourage learners from continuing their discourse. This finding can also be viewed in terms of number of utterances.

Table 2. Comparison of occurrences of teacher discourse context between the Individual context and the Individual plus Own utterance contexts

	Assertion	Expli- cation	Feed- back	Info- question	M- checking	M- seeking	Modif. interest
Individual	9	15	27	8	52	21	29
Indi.+Own Utt.	15	15	39	8	68	34	64
Progression	+67%	+0%	+44%	+0%	+31%	+62%	+121%

In Table 2, the progression of each type of discourse context refers to the percentage of utterances added after the first utterance within a single learner turn. The change after the inclusion of data from the Own utterance context indicates the multi-utterance tendency of the turns. The striking increase in the Modified interaction context means that this type of discourse context had the strongest tendency to produce more than one utterance. In other words, this context does not appear to dissuade speakers from prolonging their discourse after rapidly finishing the topic at hand. Thus, the “lateral sequence” (Jefferson 1972) in classroom settings may open and close very rapidly. In the same way, Assertion and Message seeking (and, to some extent, Feedback) appear to exhibit the same tendency. It is therefore clear that Modified interaction is a context that does not immediately give rise to discourse complexity in terms of markers, length, or topicfulness but can render a discourse complex due to discourse continuity. Message seeking and Feedback, then, may be the contexts best suited to creating discourse complexity and utterance generation.

#### 5.4. Summary

The data gathered above suggest that the pragmatic and interactional context which a teacher creates through his or her discourse plays a role in the qualitative and quantitative characterization of learner discourse. When they produce discourse within their own turn, or when the teachers support their talk with short feedback, learners tend to produce longer, more marked and more topicful discourse than in over-regulating contexts. Likewise, learner discourse varies across interactional contexts. In their exchanges with the teacher, learners construct their utterances with more discourse markers and more complex discourse than when they interact

with their classmates. Moreover, certain types of teacher discourse act as subtle catalysers for increasing the complexity of the learner's discourse. Although the manifestation of subjective modulation appears to be less context-sensitive, the use of discourse connectors, utterance length, and topic development seem to largely depend on the type of pragmatic context. As a whole, a moderately regulating context (here, Message seeking) seems more suited to encouraging longer and richer speaking turns than does a free or more constraining context (here, Assertion, Explication, and Message checking). A difference in the pragmatic discourse context also seems to be linked to discourse continuity. Whilst Modified interaction was the most limiting and did not make the next utterance complex in any respect, it did not discourage learners to pursue their talking by producing a multi-utterance discourse. Finally, this study shows that it is possible to look into the impact of discourse and interactional context on discourse complexity by taking into account a single utterance as well as a chain of utterances within a turn.

## **6. Study II: Context analysis of classroom interaction**

The overall analysis of discourse presented above provided general insights into the characteristics of the learner discourse in relation to the interactional and pragmatic contexts of the teacher's discourse. However, this discourse analysis was not intended to describe each instance of the teacher/learner interaction. The present section will illustrate what was induced from the above analysis and also describe other aspects of classroom interaction which were not taken into account. For this reason, some of the examples selected below might appear contradictory to the findings of the discourse analysis. The examples are aimed at deepening the above findings by bringing multiple facets of classroom interaction to light.

### **6.1. Interactive regulation of discourse continuity and coherence**

In the last section, it was assumed that learners take advantage of some of the interactional and pragmatic contexts created by the teacher's discourse to make their discourse multi-utterance. However, it seems that the multi-utterance tendency is qualitatively variable. To illustrate this, three examples will be compared.

First, take the example where a Vietnamese learner (SV) is talking about his country. The opening context of this example is Message seeking proposed by the teacher (T), which is likely to render the discourse long, highly marked, topicful, and multi-utterance. (The English translation of each example and the transcription conventions are given in the Appendices).

- (1) 234T : bon les autres\ + les autres alors\ \ +++ SHOh::\ \ +++ ce dont tu es  
 235SV : 

plus fier\
mh:::\

 ++ moi j'ai:::\ ++ c'est vraiment énorme\ \ ++ depuis je suis en France\ + je:: veux::\ ++ je connais\ ++ beaucoup beaucoup de choses\ ++ que je:: ne::\ + je n'ai jamais imaginé\ \ +++ chez moi\ \ +++ peut-être j'ai déjà vu dans la télévision\ + j'ai:: vu::  
 → 236T : à la télévision\ +  
 → 237SV : à la télévision\ + je:: + entendu dire ++ n:: m'attends comme ça::// ++ mais:: je ne peux pas + imaginer chaque moment comme ça\ \ +++ par exemple\ \ +++ (j'ai xxxxx) comme compensation\ ++ euh:: pour les français qui ne travaillent pas euh:\ + qui:: n'a pas:: la capacité ++ de:: travailler\ \ +++ c'est:: xxxxx une compensation\ \

SV is talking about the topic at hand by involving his personal feelings via a comparison with his home country, and his discourse is modulated. He starts his talk (235) with a short preliminary utterance following an opinion-pronoun modulator *moi* ('me'). From this, a relatively abstract argument is advanced with a repetitive modulation (*beaucoup beaucoup de* 'lots lots of') which is temporarily contextualized (*depuis* 'since'). SV develops this argument with a comparative association between France and his own experiences of his home country. During this time, the teacher (T) comments upon a small linguistic problem, which SV reflects on only briefly. That is, T's intervention does not, as mentioned above, dissuade SV from continuing his discourse. After that, he exemplifies the topic discussed so far, thereby enriching his discourse. His argumentative approach is progressive with the appropriate use of markers, without being a boring topical progression.

The preceding example demonstrates the principle of communicative motivation by the teacher's discourse noted in the discourse analysis above, i.e. the learner's discourse is rendered complex by way of a moderately regulating context generated by the teacher's discourse. The next example of a Spanish learner (LE) suggests another aspect of this principle. This conversational sequence, which starts in a free, collective context, is LE's reply to a previous exchange about another learner's remark regarding the absence of patriotism in France.

- (2) 496LE : moi/ + je trouve ++ [le] contraire\\ +  
 497T : bon\\  
 498AP : ((laughing))  
 → 499LE : par rapport à ((cough)) à l'espagne\\ + ici\\ + la france/ c'est le::\\  
 c'est\\ ++ ah::\\ d'abord on est français\\ ++ ensuite on est:: breton::/  
 on est:: provença::l/ et:: on est n'importe quoi/ mais d'abord  
 500T : 

on est
--------

 on est européen ((cough))  
 → 501LE : et (voilà) bon\\ je sais pas comment classer ça/ ++ mais:: n:: +  
 parce que l'espagne/ c'est::\\ c'est [le] contraire\\  
 502T : hum hum/  
 503LE : tu es (race)\\ 

tu es catalan\\
-----------------

  
 504YJ : 

ah:::::
---------

  
 505LE : tu es:: d'un::: 

xxxxx ((head shake))
----------------------

  
 506T : 

ah::\\ ça c'est intéressant\\
-------------------------------

  
 → 507LE : être espagnol/ bon::\\ c'est:: ((head shake)) pas toujours bien\\ fu-fu-  
 fu ++

The first utterance (496) is a preliminary remark which already gives the argumentative direction of her upcoming discourse. This utterance is clear, by its virtue of briefness and the presence of the opinion pronoun *moi* ('me'). After T's backchannel signal (497) – a discourse-generating context – and the laughter of all participants (AP) (498), she introduces a comparative clause and hints at her conclusion with a consequential exemplification (499). Just when she returns to her main claim with *mais* ('but'), T attempts to pursue the previous development of LE's discourse (500). LE dodges this proposal rapidly, casting it aside (501), and introduces her antithesis about Spain with *mais* ('but'). T responds with a backchannel signal (502). Then, the same exemplification is given in order to contrast the opposite case (503). Finally, LE's pre-announced claim is *de facto* proven, and she modulates the ending by a meaningful head shake and laugh.

Whilst this learner's discourse can be regarded as good argumentation, it is not linked to discourse complexity within a turn but through several turns. Indeed, the context which the teacher created (feedback and assertion) is a multi-utterance generator, as shown in Table 2. This might be another type of moderate regulation of discourse, more or less controlled by the teacher, who attempts to get engaged in the interaction with the learner by contributing, in collaboration with her, to a paced and argumentative development of the learner's discourse. In other words, a context that allows learners to produce a complex, topicful, marked, and/or coherent discourse can be co-constructed by the teacher and learners. This suggests that this kind of dis-

course complexity can be generated by the pragmatic regulation of the immediate context of discourse production (as discussed above in Section 5) as well as by the interactional regulation of the learner's discourse through a sequence of turns.

Moreover, it could be argued that too free a context might be linked to a meaningless, long discourse and hence, one that is less advantageous for the discourse learning. Consider the case of a Chilean learner (MC). The sequence starts with Message seeking, which gives MC the right to speak freely.

- (3) 414T : tu voulais dire quelque chose//  
 415MC : oui\ ah::\ ++ moi j'ai une autre impression dans:: un français\ + dans:: les bureaux// + non\  
 416T : dans les bureaux// + dans:: l'admini [stration\]  
 417MC : [ dans ] l'administration\ + ah:: dans mon expérience/ c'est:: LÀ/ ++ c'est là/ + ou:i\ où j'ai:: été + bien traitée/ + [ par ]  
 418T : [ c'est ] là où j'ai été [ bien traitée\ ] ++  
 → 419MC : [ oui\ ] ++  
 MAIS\ + je CROIS + [ki:: SA]\ ((=?)) + ah::\ ++ (Ç) peut-être pour:: + (être très)\ + parce que:: + [il sa::]\ ++ [il son] ehm ((start-up gesture))\ [n sui] [n suis]\ [n sui] MOI// ++ non non\ qui est dans le:: [enfon de:: u]\ + sinon [il mon di]// ((reported discourse)) où tu viens\  
 420T : ah:: attends attends attends\ + parce que + ce:: ce:: ((...))

MC begins to present her opinion about the French (415) and her experience with them (417), with in each turn including some corrective feedback from the teacher. After that, she introduces her main claim (419). This turn is full of discourse markers and utterances. Yet, her message is difficult to understand, since the use of discourse markers is inappropriate and lacks logical consistency. Presumably, MC's attention is too actively focused on message production, thanks to the contextual freeness that T allows her, but it does not cover her discourse coherence or cohesiveness. Note that this learner was the most active participant in the group, and her discourse tended to be wordy and prolix. She looks like one of the "hyperactive" learner types reported by Allwright (1980). MC's discourse, then, should have been regulated more by the teacher so as to make her claim clearer, if possible<sup>6</sup>. This example suggests that appropriate and coherent learner discourse might be generated through and by the interaction process.

## 6.2. Multi-utterance discourse and turn management

It sometimes happens that a context aimed at promoting multi-utterance performance by the learner fails to work due to the teacher's turn management. Consider this example of a Japanese Learner (YJ).

- (4) 040T : ((deictic gesture)) vous avez la même [différence + en japonais//]  
 041YJ : euh:.....\  
       déjà c'est différent\ + an et année\ ++  
 042T : ah\ [ vous avez la même différence\ ]  
 043YJ : [ le mot est différent\ ] ++  
       oui le mot est différent\ +  
 → 044T : oui\  
 045YJ : déjà ++ c'est pas le même euh:: ++  
 → 046T : Aline\ ++ c'est pareil// ++

Here, the Message checking context (040) comes up after a question/response exchange. Although this type of discourse, as mentioned above, is less likely to encourage complex or continuous discourse, this time it has seemingly enhanced YJ's desire to develop the ongoing topic, since after her first attempt, unsuccessful due to overlap (043), a resumption is observed by the end of T's utterance (042). Therefore, YJ's discourse in turn (045) is likely to take place within a multi-utterance movement. Her objective, however, is not attained by T's allocation of the turn to another learner (046). It is difficult to know whether this was caused by YJ's long hesitation or by T's hasty turn management.

The next example is a similar multi-utterance failure with a Cypriot Greek-speaking learner (VG). The conversational topic is the same as that in the previous example.

- (5) 056T : et et toi/ Vassiliki\  
 057VG : quel- quelquefois oui/ quelquefois non\  
 → 058T : d'accord\  
 059VG : par exemple\ on dit je reste une année/ mais les années soixante  
       c'est:: +  
 060YJ : ((head shake)) oui oui\  
 061VG : c'est pas le::  
 → 062T : oui ça c'est intéressant cet exemple\ les années soixante\  
 063YJ : les années soixante\  
 064T : les années quatre-vingt-dix d'accord\ ++ euh:: donc::: c'est ça\ +  
       une + d'autres questions// ++



This sequence starts with a relatively opaque Message seeking context, but situational information is clear enough for VG to take in the hidden message. T's feedback (058) then encourages VG to continue her discourse. Meanwhile, it turns out that the topic that VG is dealing with is at once picked up as interesting by T (062), who goes on to the next topic (064). Such an interruption must leave the learner with a feeling of incompleteness. In both the last two examples, the interactional management of the discourse was problematic: although the teacher's discourse opened up a multi-utterance context, she threw away a potentially good opportunity for the learner to produce continuous discourse. Thus, the teacher seems to have failed to manage discourse continuity context.

However, it was not always the teacher who misled the discourse context. Consider the next example. The first part of this sequence is a classic classroom exchange: the teacher starts with an utterance whereby she seeks participation from the learners. Here, a Korean learner (ACo) responds to this request (087).

- (6) 088T : c'est quoi + DONT vous avez peur\\ +  
 087ACo : le temps\\ +  
 088T : les temps\\ ++  
 089ACo : par exemple\\ +++ je ne peux pas + distinguer entre passé + passé composé + l'imparfait et plus-que-parfait\\ ++ si + tout est mélangé/ + je ne comprends pas\\ ((laughs)) +  
 → 090T : est-ce que\\ + est-ce que tu as compris quelque chose là-dedans\\ ++ passé composé ++ imparfait ++ plus-que-parfait\\ ++  
 091ACo : oui\\  
 → 092T : est-ce que tu as quelques idées/ + oui/ + c'est sûr\\  
 093ACo : oui oui\\  
 094T : alors quand tu dis/ + je je mélange tout\\ + je crois que tu es perdue:/ mais +++ peut-être pas tout à fait\\ +  
 095ACo : he-he-he-he ++

ACo develops her claim in her second utterance (089), which is sufficiently long and marked. Nevertheless, the second part is not like the first: ACo's performance is not rich in discourse complexity (090–095). Perhaps her reluctance to speak comes from a misinterpretation of the pragmatic device (090). From a formal viewpoint, *Est-ce que tu as compris quelque chose là-dedans?* ('Have you understood anything in that?') is a Yes/No question, but in terms of function, this utterance means *Qu'est-ce que tu as compris là-dedans?* ('What have you understood in that?'), which is a discourse-rich Message seeking context. Admittedly, a question with an indefinite

pronoun (e.g. *quelque chose* ‘anything’) can often be pragmatically interpreted as a *wh*-question, as in this example. Now, ACo, who only minimally responds to T having taken her utterance to be a simple question, has not appropriately interpreted the illocutionary meaning of T’s utterance, which attempted to elicit a more complex reaction. It is possible that such pragmatic ambiguity is what made the exchange less topicful.

### 6.3. Proficiency and multi-utterance discourse

The last example also suggests that multi-utterance production might need minimal pragmatic knowledge and hence depends on the learner’s discourse competence. Let us analyse the performance of two Korean learners. They produced fewer continuous utterances than all other participants. Their discourse never exceeded three utterances within a single turn, and its structure was often elementary. The next example starts an open context, with a topic previously tackled by other learners (*viz.* qualifiers used to describe the French people).

- (7) 522ACo : pour moi/ ah::\ ++ à l’extérieur/ + d’habitude les français sont très gentils\ ++ par exemple\ [il di] toujours\ bonjour merci au revoir comme ça:/  
 523T : ça c’est gentil ou c’est [ko-] euh:\ POLi\ euh ils sont POLis\ +  
 → 524ACo : ah: oui oui\ + mais quand j’étais:: ++ en bel- + en belgique\ + [il n di] pas toujours bonjour merci au revoir comme ça\ ++ alors:: n:: +

In the first utterance, ACo defends her opinion, but receives a reply from T (523) since there is, at least for T, a flaw in ACo’s argument. Although the learner tries to make a comparison as a counter-argument in order to refute T’s reply, her argumentative structure with *mais* (‘but’) appears not to be powerful enough to be convincing. In addition, the incomplete discourse she voluntarily uses states her claim only implicitly, and such rhetoric is not efficient as an argumentative device in this context. The use of utterances by another Korean learner (JCo) is not any better.

- (8) 793T : ((head nod)) +++ julie\ ++  
 794JCo : ++ n:: ha-h ++  
 795T : les français sont timides\ + c’est vrai//  
 796JCo : ah-HA-HA-HA  
 797SV : pas du tout\ +  
 → 798JCo : je sais pas::\ H-ha-ha-ha ++++ n:: + n:: ++ je me sens qu’ils sont gentils\ ++ oui::\ +++ je(h) sais(h) pas(h)\ h-h-h-h

The context which the teacher proposes is Message checking (795). JCo's discourse consists of three utterances not linked by any connector. Two of the utterances are merely simple expressions of modulation. Furthermore, JCo's main claim (798) is only the resumption of an opinion previously used by another learner (not described here). In fact, it might not be the teacher's elicitation technique that is problematic, but the learner's discourse competence itself. A review of all of the longest utterances of the two Korean learners during the class revealed that their discourse was mostly made up of only two arguments linked by an elementary connector (e.g. *mais* 'but', *par exemple* 'for example', *si* 'if', *quand* 'when') or not linked at all. Therefore, in the case of these learners, it would be difficult for a supposedly discourse-rich or multi-utterance context to function as expected. This suggests that multi-utterance performance requires minimal knowledge of argumentative structures and markers.

This issue could be also interpreted in another way: learners are merely reluctant to produce multiple utterances, so their discourse structure or use of markers is elementary. Admittedly, it has been often pointed out that some types of learners are not inclined to participate in classroom interaction, especially Asian learners<sup>7</sup>. At the present time, no data or methods are available to further examine this issue. We will return to this question below.

## 7. Discussion

This study suggests that a two-fold analysis can be fruitful for understanding teacher/learner interaction in the classroom. Some findings of the discourse analysis and the context analysis were convergent, whilst others revealed different aspects of teacher/learner interactions in a conversational classroom.

Let us now return to the issues posed in Section 4, namely: 1) the effect of teacher discourse on learner discourse; 2) the possibility of triggering discourse learning in the conversational classroom. Note that the statements formulated below must be interpreted with caution, since the data used in this study were obtained from only one hour of recording. Moreover, no longitudinal data were examined. This means that some of our statements will need further examination in order to determine real acquisitional effects.

Concerning the first issue, it turned out that there are some principles governing the teacher's discourse that can motivate learners to produce

extended discourse. The principle of moderate regulation of context was found to be a teaching device that motivates learners to produce discourse. This principle can be applied not only to the pragmatic and/or interactional choice of the teacher's discourse for the next turn, but also to the more subtle regulation of the interaction process in order to create a multi-utterance context through a sequence of turns. Excessive freeness or control of the pragmatic or interactional context generated by the teacher's discourse appears to be unfavourable to the learners' production of complex or coherent discourse. Moreover, it seems that teachers should pay careful attention to the discourse context, or else a potential multi-utterance context might be spoiled by awkward management of the interaction. This points out how very difficult it is to handle interactions in a conversation-based classroom. Nevertheless, these classroom principles are probably the same in interactions outside the classroom.

Other discourse-motivation principles are applicable to the second question, since they imply some of specificities of classroom interaction. These principles turned out to be disappointing for teachers, however: the learners' discourse was better when they produced utterances within their own speaking turn and when the teacher supported their discourse only with feedback. This could mean that learners' best discourse performance does not need a trained teacher after all but only a good listener. In other words, discourse learning or acquisition may not require an instructional settings.

Based on comparisons of the prior-speaker's utterances, it turned out that the learners' discourse was marked half as much and half as long when the prior speaker was a classmate than when the person was the teacher. Note also the following qualitative difference in discourse between the learner and teacher contexts: the learners used more feedback, questions, replies, and turn initiations with their classmates than with their teacher (see Figures 1 and 2). This indicates that when learners are talking with a classmate, they tend to produce a discourse closer to that of spontaneous conversation rather than generating elaborate discourse. Thus, it is assumed that whilst interactions between learners are adequate for improving "interactive competence" (see He & Young 1998: 6–8)<sup>8</sup>, interactions with a teacher enhance learners' experience of more complex discourse. Therefore, classroom interaction probably provides a productive basis for enhancing discourse competence. This finding is consistent with the study by Varonis and Gass (1985) on the negotiation of meaning in a dyadic context: learners more frequently negotiate meaning – by way of less elaborate discourse – when their partner is a non-native speaker than when he or she is a native speaker. This, however, is still not absolute evidence of the advantage of

teacher discourse and instructional settings for learner discourse acquisition. No data on learner discourse outside the classroom are available for direct comparison to classroom interaction data. But as far as the comparison between the teacher and learner contexts is concerned, and insofar the learners relate their “unhappy” discourse life outside the classroom, they would certainly benefit from interaction with the teacher that motivates the production of elaborate discourse. This issue needs further investigation to verify the actual effectiveness of classroom interaction for discourse learning.

Another specificity of classroom interaction concerns the compatibility of discourse learning and linguistic learning. It turned out that the teacher’s linguistic interventions did not prevent the learners from achieving multi-utterance performance. If this is true, teachers do not have to be afraid of dealing with learners’ linguistic problems even if it does not directly concern discourse learning. This could imply that, with regard to linguistic material, there is a specific relationship between the learners and the teacher, a sort of tacit “didactic contract” (Faraco and Kida 1999), and “face-work” (Goffman 1955) in the classroom functions in a different fashion than in a non-instructional environment.

It seems that this matter has not been pointed out enough in the second language acquisition literature in models proposed for linguistic acquisition such as “conversational adjustment” (e.g. Long 1996), “negotiation of meaning” (e.g. Varonis and Gass 1985), “sequence of potential acquisition” (De Pietro, Mathey, and Py 1989), “collaborative dialogue” (Swain 2000)<sup>9</sup>. These models have focused primarily on linguistic aspects of the acquisitional activity in the classroom, but have not sufficiently discussed the potentials of discourse acquisition (but see Long 1996). Indeed, negotiated adjustments of the teacher’s discourse result from “a part of the process of trying to communicate with learners of limited competence” (Ellis 1994: 257) occurring normally and spontaneously in the classroom and therefore have a legitimate place in the second language class. However, it seems possible to account for the classroom acquisition of discourse by second language learners even within these models of linguistic acquisition. Accordingly, one should reconsider classroom-based studies from the standpoint of the process of learner discourse acquisition, because classroom interactions are not limited to grammar or vocabulary learning but can offer learners opportunities for learning the ability to conduct argumentative or narrative discourse.

It was also shown here that the effectiveness of teacher discourse improving learner discourse might have its own limits for some learners. When discussing the Korean learners, it was suggested that their multi-utterance

behaviour might depend either on their proficiency or on their culture-specific communicative behaviour. Probably, the self-confidence learners have of their communicative proficiency also plays a role in their performance. In any case, this kind of learner continues to produce short discourse so their discourse acquisition remains at risk. To find out whether explicit discourse learning would improve their performance and to understand the real cause of their reluctant performance, further investigation of the second language classroom is needed.

## **9. Conclusion**

This chapter focused on teacher/learner interaction in a conversational second language classroom in an attempt to find out more about the potentials of discourse acquisition in instructed contexts. The teacher's discourse in a classroom can have positive effects on the learners' discourse in such a way that the learners produce meaningful and complex discourse that promotes discourse acquisition. However, there are some cases where the teacher's discourse and interaction management is little efficient and it is necessary in these cases to clarify the reasons for reluctant discourse production. Since discourse learning is not necessarily incompatible with linguistic learning, it is possible to orient the classroom-based study of second language acquisition toward the search for models of classroom interaction which are a rich potential for discourse acquisition.

## **Notes**

1. I would like to express my gratitude to Alain Giacomi, Michel Paradis, Martine Faraco, Jane Orton, Michele de Courcy, Vivian Waltz and anonymous reviewers for their helpful comments on earlier versions, and to Collen Masson and Sarah Tatz for their invaluable assistance. Nevertheless, I alone am responsible for any problems that remain. I also gratefully acknowledge the encouragement of Alex Housen.
2. In this chapter, I use the term "acquisition", knowing that the expression "acquisition of discourse competence" would, in my view, be inadequate, insofar as each learner already has a discourse competence in his or her native language.

3. According to Paradis (1998), the working mechanism of pragmatic and discourse competence (relative to knowledge of discourse structures and non-literal meanings) is particularly vulnerable to right hemisphere damage, whereas traditional linguistic and metalinguistic knowledge (syntax, morphology, phonology, and lexicon) are also to left hemisphere damage.
4. Declarative memory subserves knowledge of a metalinguistic, episodic, or encyclopaedic nature, which can only be learned consciously (i.e. by noticing the items to be learned) and used in a controlled manner. Procedural memory supports a wide range of cognitive and sensorimotor activities governed by action programs and subserves its language related part, called “implicit linguistic competence” (Paradis 1994), which is acquired incidentally and employed automatically.
5. Interestingly, it has been shown that the occurrence of repetition does not vary according to the linguistic proficiency level of the L2 classroom (Faraco, 2002).
6. Such classroom management is not always easy, however, as Allwright (1980: 182) describes: “The teacher takes the discourse maintenance task as Igor stops, and faces a problem now. Strictly speaking Igor has left no task except that at the level of turn taking (the problem of leaving someone else to maintain the discourse) but a basic “rule” of discourse is that successive turns shall be topically relevant to each other. To reject it might suggest an unwillingness to allow real communication to flourish in the classroom, whilst to accept it means a disruption of the lesson plan, which might be resented by other learners and might even lessen their confidence in her ability to control events. In fact, the teacher accepts the change of topic category and sets a further task for Igor by suggesting how his contribution might fit into the general argument, and requesting confirmation from Igor.”
7. According to Ellis (2000), “[i]t is possible, however, that for some learners at least the act of trying to produce the L2 in the public forum of the classroom is so stressful as to distract from their ability to process input. [...] Japanese students are well known for their reluctance” (p. 246). The same thing can be said of turn-taking behaviour (Markel 1975).
8. Interactive competence refers to the aptitude for metadiscursive management of turns and topics and for signalling boundaries, and to several strategic means used in local interactive practice, especially in spontaneous conversation.
9. This concept, defined as “[...] dialogue in which speakers are engaged in problem solving and knowledge building” (Swain 2000: 102), is an application of the Vygotskian concept of social interaction for conceptual acquisition (see Wertsch e.g. 1991). The structure of talk that this dialogue refers to, in my view, is that of spontaneous conversational interaction.



## References

- Allwright, R. L.  
1980 Turns, topics, and tasks: Patterns of participation in language learning and teaching. In *Discourse Analysis in Second Language Research*, D. Larsen-Freeman (ed.), 165–187. Rowley, MA: Newbury House.
- Cohen, N.  
1984 Preserved learning capacity in amnesia: Evidence for multiple memory systems. In *Neuropsychology of Memory*, L. R. Squire and N. Butters (eds.), 83–103. New York: Guilford Press.
- De Pietro, J.-F., M. Matthey and B. Py  
1989 Acquisition et contrat didactique: les séquences potentiellement acquisitionnelles dans la communication exolingue. In *Actes du 3e colloque régional de linguistique*, D. Weil and H. Fugier (eds.), 99–124. Strasbourg, FR: Université des Sciences Humaines et Université de Louis Pasteur.
- Ellis, R.  
1994 *The Study of Second Language Acquisition*. Oxford, UK: Oxford University Press.  
2000 *Learning a Second Language Through Interaction*. Amsterdam/Philadelphia: John Benjamins.
- Faraco, M. and T. Kida  
1999 Acquisition de L2 et séquence d'apprentissage: degré de stabilité interactive. *International Review of Applied Linguistics in Language Teaching* 37 (3): 215–230.  
forthc. Some notes on gesture in the second language classroom. In: S. G. McCafferty and G. Stam (eds.), *Gesture in Second Language Acquisition*. Laurence Erlbaum Associates.
- Faraco, M.  
1995 La répétition dans la communication exolingue en classe de FLE. *International Review of Applied Linguistics in Language Teaching* 33 (3): 217–250.  
2002 Répétition, acquisition et gestion de l'interaction en classe de L2. *Acquisition et Interaction en Langue Étrangère* 16: 97–120.
- Goffman, E.  
1955 On face-work: An analysis of ritual elements in social interaction. *Psychiatry* 18: 213–231 (Reprinted: J. Laver and S. Hutcheson (eds), *Communication in face to face interaction. Selected readings*, 319–346. Harmondsworth, UK: Penguin Books.
- He, W. A. and R. Young  
1998 Language proficiency interviews: A discourse approach. In *Talking and Testing. Discourse Approaches to the Assessment of Oral Proficiency*, R. Young and W. A. He (eds.), 1–24. Amsterdam/Philadelphia: John Benjamins.



- House, J.  
1996 Developing pragmatic fluency in English as a foreign language: Routines and metapragmatic awareness. *Studies in Second Language Acquisition* 18 (2): 225–252.
- Hymes, D. H.  
1984 *Vers la Compétence de Communication*. Paris: Crédif-Hatier (unedited version in English).
- Jefferson, G.  
1972 Side sequences. In *Studies in Social Interaction*, D. N. Sudnow (ed.), 294–338. New York: Free Press.
- Markel, N. N.  
1975 Coverbal behavior associated with conversational turns. In *Organization of Behavior in Face-to-Face Interaction*, A. Kendon, R. M. Harris and M. R. Key (eds.), 189–197. The Hague: Mouton.
- Long, M. H.  
1996 The role of the linguistic environment in second language acquisition. In *Handbook of Second Language Acquisition*, W. C. Ritchie and T. K. Bhatia (eds.), 413–468. San Diego: Academic Press.
- Long, M. H. and C. S. Sato  
1983 Classroom foreigner talk discourse: Forms and functions of teachers' questions. In *Classroom-Oriented Research in Second Language Acquisition*, H. B. Seliger and M. Long (eds.), 268–285. Rowley, MA: Newbury House.
- Lörcher, W.  
1986 Conversational structures in the foreign language classroom. In *Learning, Teaching and Communication in the Foreign Language Classroom*, G. Kasper (ed.), 11–22. Aarhus, DA: Aarhus University Press.
- Kendon, A.  
1990 Some context for Context Analysis: A view of the origins of structural studies of face-to-face interaction. In *Conducting Interaction. Patterns of Behavior in Focused Encounters*, A. Kendon, 15–50. Cambridge, UK: Cambridge University Press.
- Kida, T.  
forthc. Does gesture aid discourse comprehension in L2? In *Gesture in Second Language Acquisition*. S. G. McCafferty and G. Stam (eds.), Laurence Erlbaum Associates.
- Moder, C. L. and G. B. Halleck  
1998 Framing the language proficiency interview as a speech event: Native and non-native speakers' questions. In *Talking and Testing. Discourse Approaches to the Assessment of Oral Proficiency*, R. Young and W. A. He (eds.), 116–146. Amsterdam/Philadelphia: John Benjamins.

- Moirand, S.  
1990 *Enseigner à Communiquer en Langue Étrangère*. Paris: Hachette.
- Paradis, M.  
1994 Neurolinguistic aspects of implicit and explicit memory: Implications for bilingualism and SLA. In *Implicit and Explicit Learning of Language*, N. C. Ellis (ed.), 393–419. London: Academic Press.  
1998 The other side of language: Pragmatic competence. *Journal of Neuro-linguistics* 11: 1–10.
- Schiffrin, D.  
1994 *Approaches to Discourse Markers*. Cambridge, MA: Blackwell.
- Swain, M.  
1985 Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In *Input in Second Language Acquisition*, S. Gass and C. Madden (eds.), 235–253. Boston, MA: Heinle & Heinle Publishers.  
1995 Three functions of output in second language learning. In *Principle and Practice in Applied Linguistics. Studies in Honour of H. G. Widdowson*, G. Cook and B. Seidlhofer (eds.), 125–144. Oxford, UK: Oxford University Press.  
2000 The output hypothesis and beyond: mediating acquisition through collaborative dialogue. In *Sociocultural Theory and Second Language Learning*, J. P. Lantolf (ed.), 97–114. Oxford, Oxford University Press.
- Varonis, E. M. and S. M. Gass  
1985 Non-native/non-native conversations: A model for negotiation of meaning. *Applied Linguistics* 6 (1): 71–90.
- Wertsch, J. V.  
1991 *Voices of the Mind: A Sociocultural Approach to Mediated Action*. Cambridge, MA: Harvard University Press.
- Widdowson, H. G.  
1981 *Une Approche Communicative de l'Enseignement des Langues*. Paris: Hatier-Credif (original title: *Teaching Language as Communication*. Oxford, UK: Oxford University Press, 1978).

## Appendices

### Transcription conventions:

[aaa]	Phonetic transcription
[ ]	Overlapping speech
:	Elongated sound or syllable ( $\pm 1/4$ second)
+	Pause ( $\pm 1/4$ second)
AAA	Emphasis stress
// (\)	Terminative rising (falling) intonation
/ (\)	Continuous rising (falling) intonation
((soft))	Description by transcriber
(maybe)	Transcription doubt
XXX	Inaudible
(Ç)	Tongue noise

### Translation of Example (1)

- 234T : right the others + so the others +++ Shoh:: +++ what you are most  
 235SV : [ proud of ] ++ me I have::: ++ it's really enormous ++ since I've  
 been in France + I:: wan::t ++ I know ++ lots lots of things ++ that  
 I:: not:: + I had never imagined +++ in my country +++ maybe I  
 have already seen in the TV + I've:: seen::  
 → 236T : on TV +  
 → 237SV : on TV + I:: + heard say ++ n:: wait for me like tha::t ++ bu:::t I  
 can't + imagine every moment like that +++ for example +++ (I  
 have ) as a compensation ++ um::: for French people who don't  
 work um: + who don't:: has:: the capacity ++ to:: work +++ it's::: ( )  
 a compensation

### Translation of Example (2)

- 496LE : me + I find ++ the opposite +  
 497T : good  
 498AP : ((laughing))  
 → 499LE : with respect to ((cough)) Spain + here + France it's the::: it's ++  
 aah:: first one is French ++ then one is:: Breton:: one is:: Provença::l  
 and:: one is whatever but first [ one is ]  
 500T : [ and finally ] one is European ((cough))

- 501LE : and (there) well I don't know how to classify this ++ bu::t n:: +  
/because Spain is::: it's the opposite
- 502T : hm hm
- 503LE : one is (race) [ one is Catalan ]
- 504YJ : [ aah::: ]
- 505LE : one is:: of a::: [ xxxxx ((head shake)) ]
- 506T : [ aah:: that's interesting ]
- 507LE : being Spanish righ::t it's:: ((head shake)) not always good fu-fu-fu ++

## Translation of Example (3)

- 414T : did you want to say something
- 415MC : yes aah::: ++ me I have a different impression abou::t French people  
+ in:: offices + no
- 416T : in offices + in:: admini [ stration ]
- 417MC : [ in ] administration + aah:: in my  
experience it's::: THERE ++ it's there + ye:s where I:: was + well  
treated + [ by- ]
- 418T : [ it's ] there where I was [ well treated ]
- 419MC : [ yes ] ++  
BUT + I BELIEVE + [ki::: SA] ((=?)) + aah::: ++ (Ç) maybe fo:::r  
+ (being very) + because::: + (they is:::) ++ (they are) um ((start-up  
gesture)) (am not) (am not) (am not) ME ++ no no who is in the::  
(infront of :: them) + if not (they said to me) ((reported discourse))  
where you come
- 420T : aah:: wait wait wait + because + this:: [ this:: ]

## Translation of Example (4)

- 040T : ((deictic gesture)) do you have the same [ difference + in Japanese ]
- 041YJ : [ um::: ]
- already it's different + 'an' and 'année' ++
- 042T : aah [ you have the same difference ]
- 043YJ : [ the word is different ] ++  
yes the word is different +
- 044T : yes
- 045YJ : already ++ it's not the same um::: ++
- 046T : Aline ++ is it the same ++

Translation of Example (5)

- 056T : and and you Vassiliki  
057VG : som- sometimes yes sometimes no  
→ 058T : okay  
059VG : for example we say I stay 'une année' but 'les années soixante'  
it's::: +  
060YJ : ((head nod)) yes yes  
061VG : it's not the::  
→ 062T : yes that that's an interesting example 'les années soixante'  
063YJ : 'les années soixante'  
064T : 'les années quatre-vingt-dix' okay ++ um:: so::: it's that +  
one + other questions ++

Translation of Example (6)

- 088T : what is it + THAT you are afraid of +  
087ACo: the tense +  
088T : the tenses ++  
089ACo: for example +++ I can't + distinguish between 'passé' + 'passé  
composé' + the 'imparfait' and 'plus-que-parfait' ++ if + every-  
thing's mixed up + I don't understand ((laugh)) +  
→ 090T : have you + have you understood anything in that ++ 'passé com-  
posé' ++ 'imparfait' ++ 'plus-que-parfait' ++  
091ACo: yes  
→ 092T : do you have any ideas + yes + I'm sure  
093ACo: yes yes  
094T : so when you say + I I mix everything up + I think you are lost but  
+++ maybe not completely +  
095ACo: he-he-he-he ++

Translation of Example (7)

- 522ACo: for me aah::: ++ outside + normally the French are very nice ++ for  
example (they says) always 'bonjour' 'merci' 'au revoir' like that:  
523T : that's nice or it's [ko-] um: POLite um they are POLite +  
→ 524ACo: aah: yes yes + but when I was::: ++ in Bel- + in Belgium + (they  
doesn't) always say 'bonjour' 'merci' 'au revoir' like that ++ so::  
n:: +

Translation of Example (8)

- 793T : ((head nod)) +++ Julie ++  
794JCo : ++ n:: ha-h ++  
795T : the French are shy + it's true  
796JCo : aah-HA-HA-HA  
797SV : not at all  
→ 798JCo : I dunno:: H-ha-ha-ha +++++ n:: + n:: ++ I feel myself they are  
friendly ++ yes:: +++ I(h) du(h)nno(h) h-h-h-h

## **Comparing the effects of instructed and naturalistic L2 acquisition contexts**





## **Second language acquisition in a study abroad context:**

### **A comparative investigation of the effects of study abroad and foreign language instruction on the L2 learner's grammatical development**

*Martin Howard*

This paper investigates a possible differential effect of study abroad and foreign language instruction on the L2 learner's linguistic development in relation to a specific component of his/her grammatical competence, namely the expression of past time. The study is based on a comparative investigation of Irish advanced learners of French during study abroad and in the foreign language classroom. By drawing on variation theory to illuminate the intrinsic detail of their linguistic development, the study investigates the effect of a range of linguistic factors on the learners' variable marking of past time. Results suggest that study abroad has a more beneficial effect than instruction by bringing about increased use of the past time markers, without changing the underlying patterns of use of such markers when those linguistic factors are present in context. The findings are discussed in terms of the universality of the acquisition process in an instructed and naturalistic setting.

#### **1. Introduction**

Second language acquisition (SLA) research has traditionally categorized the second language (L2) learner's domain of acquisition in terms of a dichotomy between classroom instruction outside the target language (TL) community and naturalistic acquisition within the TL community. Freed (1991: 5) notes that such a dichotomy between so-called *foreign language learning* and *second language acquisition* is primarily a historical phenomenon, which arose out of nationalist sensitivities after the Second World War during discussions about language learning within international organisations. However, as Gass (1990) indicates, the current relevance of such a

distinction is questionable, since it ignores the fact that alternation between both environments is the norm for an increasing number of L2 learners who participate in *study abroad*, whereby the instructed learner assumes the status of the naturalistic learner during a period of residence in the TL community. Study abroad is facilitated by a number of international exchange programmes, which typically allow L2 learners to spend part of their academic programme in the TL country<sup>1</sup>. Extensive funding of such programmes, as well as increased emphasis on the incorporation of an international dimension to both language-oriented and non-language-oriented academic programmes point to the firm belief of educators in the benefits of study abroad<sup>2</sup>. It therefore must be assumed that classroom instruction, in some way, does not facilitate the learner's development to the same extent. Such benefits are typically considered to be multiple, including those relating to personal, social, intercultural and academic issues. Whilst such issues have gone relatively uninvestigated, studies reported on by Bryant (1995), Hannigan (2001) and Stephenson (1999) none the less offer an insight into the type of attitudinal and motivational changes which study abroad learners undergo, as well as changes in their professional career plans. However, whilst such findings reflect the important benefits of study abroad, one of the principal reasons for participation in study abroad is to bring about increased L2 proficiency<sup>3</sup>. Indeed, the many language learners who venture abroad each year principally do so in the belief that, by simply 'picking up' the foreign language whilst abroad, they will become 'fluent' in that language. Such a folk belief implies that the informal exposure to the L2 in the TL community following a course of formal instruction, in some way, constitutes optimal conditions for L2 acquisition.

Whilst such a folk belief may exist, it is vital that empirical research detail its validity by identifying the relative linguistic benefits of study abroad compared to instruction in the foreign language classroom. On the one hand, such research entails important practical repercussions, given the extensive international financial funding which facilitates participation in study abroad, which could be otherwise utilized in facilitating the language learning process in the classroom. For example, it is necessary to identify whether linguistic development is uniform across the various components which constitute the learner's global linguistic repertoire in the L2, or whether study abroad facilitates greater development on certain components which classroom instruction does not, and *vice versa*. Furthermore, it is also necessary to consider the question of how the gains made by the study abroad learner might be maintained on his/her return to the foreign language classroom. For example, Raffaldini (1987) reports increases in

lexical, morphological and structural errors in the year following the learner's return from the TL community, in spite of instruction in the foreign language classroom in that time.

On the other hand, the consequences of research in this area also extend to practices surrounding the incorporation of study abroad in academic programmes: it is important that programme organisers and curriculum developers become aware of the sociobiographical and circumstantial features which impinge on the ultimate success of study abroad in bringing about increased proficiency in the learner's chosen L2. For example, little is known about how those features which define the individual as a language learner such as personality, motivational and affective factors affect his/her ultimate success in learning the L2 whilst abroad. In her study of American study abroad learners in France, Freed (1990) offers an insight into how such factors might affect the learner's linguistic development. She finds an effect for the type of interaction in which the learners engage whilst abroad, as defined by passive or active interactional activities, where passive interaction concerned reading and listening activities, whereas active interaction concerned more communicative interaction with native speakers. Grammatical gains were noted in the case of those learners engaging in active activities, whilst fewer gains were observed in the learners whose interaction was more passive. However, Freed notes that the type of interaction in which the learners engage, might be related to personality, and motivational issues. Indeed, it may be the case that classroom instruction is more suitable to certain learner 'types' than others. Furthermore, findings are similarly limited with regard to circumstantial questions, such as the timing of study abroad within a programme of instruction, length of programme, type of residence and the learner's 'raison d'être' whilst abroad, as well as the benefits of instruction whilst abroad<sup>4</sup>.

The questions outlined point to the extensive range of issues surrounding 'study abroad' within SLA research, such that a comprehensive investigation of 'study abroad' is necessarily quite expansive<sup>5</sup>. It is therefore beyond the scope of this paper to detail the large range of issues which impinge on the ultimate linguistic success of the study abroad learner, as opposed to the purely instructed learner. This paper will therefore be restricted in focus, on the one hand, to the issue of the specificity of linguistic development facilitated by study abroad as opposed to foreign language instruction. In particular, the paper reports on the relative effect of study abroad as opposed to foreign language instruction on development of the L2 learner's grammatical skills. An investigation of such issues is particularly important within the field of SLA, since, as noted by Coleman (1997), it can provide

an insight into the potential universality of the L2 acquisition process across different learning contexts. However, it is important to bear in mind that the study abroad learner is not wholly a naturalistic learner, such that the comparison is between the 'instructed and naturalistic' learner, on the one hand, and the purely 'instructed' learner, on the other.

## **2. Literature review**

As noted by Freed (1998), we still know relatively little about the intrinsic changes which characterize grammatical development in a study abroad context, as opposed to in the foreign language classroom. Indeed, grammatical development has been principally investigated in terms of differences in the level of structural accuracy attained by the study abroad learner and the instructed learner, rather than providing a more in-depth picture of the underlying differences that may characterise how the study abroad learner and the purely instructed learner differ in their use of the TL grammar.

For example, Freed et al. (1998) investigate the grammatical skills of their American learners of French participating in a study abroad programme in France and instructed learners in the US. The study was based on a comparative analysis of the learners' written productions. In terms of grammatical accuracy and syntactic complexity, no differences were noted between the learners. Furthermore, results based on native speaker reactions to the written texts also failed to identify any differences.

Similarly, Huebner (1995), who investigates grammatical development in the early stages of acquisition, fails to find differences between his American learners of Japanese in Japan and instructed learners of Japanese in the US. He presents a comparative investigation of their grammatical development in L2 Japanese. His results suggest little difference between the two groups of learners. As such, Huebner notes that, whilst study abroad is not any more beneficial than classroom instruction, neither is foreign language instruction any more beneficial than study abroad. Furthermore, given that the learners were in the early stages of acquisition, the results further point to the fact that study abroad can be equally beneficial at this level of proficiency as foreign language instruction. Such a finding contrasts with the general belief that beginner learners are at a linguistic disadvantage which prevents them from reaping the benefits of study abroad in the same way as more proficient learners.

A final study which fails to find a superior effect for study abroad on the L2 learner's grammatical accuracy is reported on by DeKeyser (1991).

He provides a comparative investigation of American learners of Spanish spending a semester in Spain and instructed learners in the US. Results offer no evidence of a more positive effect for study abroad on the learners' grammatical accuracy. This is in spite of the fact that the study abroad learners made more significant gains in fluency and lexical development.

In those studies where a more positive effect has emerged for study abroad than instruction, an effect for certain linguistic factors has generally been noted. For example, Freed (1990) finds that the grammatical gains made by her American learners of French spending six weeks in France differed as a function of their proficiency level. Results of a grammatical test administered before and after residence abroad suggest grammatical gains for the less proficient learners, whereas no development was detected for the advanced learners. Unfortunately, the nature of the test instrument administered which simply provides a single test score does not specify the nature of the grammatical gains made by the learners.

Whilst Freed notes an effect for the learner's proficiency level, Möhle and Raupach (1984) also propose an effect for the target language in their crosslinguistic investigation of French learners of German in Germany and German learners of French in France. Whilst the learners of French demonstrate no grammatical gains, development was apparent in the learners of German, as noted through increases in grammatical accuracy and syntactic complexity. The authors explain the discrepancies in results in terms of an effect for the German grammatical system: development was particularly noted on case endings, which are an integral component of German grammar. However, such a conclusion is rather puzzling, given that French has a relatively rich grammatical morphological system as well, albeit not for case.

Whilst development on the learner's grammatical skills during study abroad has primarily been investigated solely in terms of changes in level of accuracy, few studies have compared grammatical development across study abroad and instructed learners in relation to their acquisition of particular conceptual entities, such as temporal and spatial reference, gender, agreement, subordination, etc. As noted earlier, however, such research is particularly important within SLA research, as it can reveal possible similarities and differences between the instructed learner and the study abroad learner in terms of their acquisition of particular structural elements. Such a more detailed insight into the learner's grammar is, however, provided by Ryan and Lafford (1992) and Guntermann (1992, 1995). Both studies consider the sequence of acquisition of the functions expressed by the Spanish copula *ser* and *estar* by American learners in a hispanophone country. A comparison is made with VanPatten's (1987) findings in the case of his

instructed learners in the US. Results point to general similarities between the sequence of grammatical development in both a study abroad context and during classroom instruction.

In summary, the studies reviewed point to the relatively similar effect for study abroad as instruction on the learner's grammatical skills, both in terms of the level of structural accuracy attained, as well as the sequence of development on use of specific L2 forms. This is true in the case of studies which measured grammatical development through formal tests which favour the deployment of explicit knowledge, but also through spontaneous communicative tasks which are more likely to involve implicit grammatical knowledge only. On the basis of such evidence, Freed (1998: 50) states, with regard to grammatical development, that "significant changes do not take place in the study abroad context". Whilst such a conclusion points to the limited effect of study abroad on the learner's grammatical development, it equally points to the fact that classroom instruction is just as successful in facilitating the learner's grammatical development. This is in spite of the restrictions which characterise the classroom as a domain of acquisition, in terms of the level of intensity of contact with the L2 and the type of contact feasible. On the other hand, however, Freed's conclusion is somewhat surprising, given that research on other aspects of the learner's linguistic repertoire in the L2 has noted important differences between the study abroad learner and the instructed learner. In particular, such differences relate to the more increased development on the study abroad learner's sociolinguistic competence, fluency, and lexical competence.

In relation to the learner's sociolinguistic competence, Regan (1998) notes that study abroad has been found to lead to a dramatic increase in use of sociolinguistic markers across typologically very different L2s, such as French and Japanese. Such a conclusion is specifically exemplified by Regan's (1995) investigation of the variable deletion of the negative particle *ne* by her Irish advanced learners of French spending a year in France. The variable deletion of this marker functions as a sociolinguistic marker in TL French, and is constrained by a range of (extra)linguistic factors. The results of the study point to the increased deletion of this marker by the learners, such that the learners' sojourn in France sensitized the learners to the variable use of this marker in a way that classroom instruction had not. Similar findings have also been noted in the case of other L2s such as Japanese. For example, Marriot (1995) investigates development on the highly complex usage of honorifics in TL Japanese by her Australian learners. In particular, she notes the learners' increased sensitivity to the factors constraining their usage.

Whilst study abroad has been found to have a highly beneficial effect on the learner's sociolinguistic competence, benefits have also been noted in relation to the learner's fluency. For example, in her comparative study of American learners of French spending a semester in France and instructed learners of French in the US, Freed (1995b) notes that such increased fluency emerges across multiple features of their speech. For example, in contrast with the instructed learners, she finds that the study abroad learners speak more and significantly faster. Their speech is also characterised by fewer dysfluencies and longer streams of continuous speech. Similar findings are reported by Lafford (1995) in her comparative study of American study abroad learners in a Spanish-speaking country and a control group of instructed learners in the US. In the related area of communicative strategies, Lafford also notes that the study abroad learners develop a more expansive range of communicative strategies, such that they are more successful at coping, linguistically, with the demands of interactive communication in real time.

Finally, the more beneficial effect of study abroad has been noted in relation to the L2 learner's lexical acquisition. For example, in their study of British study abroad learners of Spanish in Spain, Ife et al. (2000) find that, not only did their learners demonstrate a more expansive lexical range following study abroad, but they also demonstrated a more native-like organisation of their lexicon. In a study which was specifically concerned with the L2 learner's verb lexicon, Howard (2002b) also presents evidence of the important lexical gains made by the study abroad learner as opposed to the instructed learner. The study was based on a comparison of Irish learners of French in a study abroad context and in the foreign language classroom. Results suggested that not only did the study abroad learners demonstrate a more expansive lexical verb repertoire, but they also acquired a more sophisticated verb repertoire.

In summary, the highly beneficial effects of study abroad in L2 acquisition have emerged in terms of the study abroad learner's increased fluency, as well as sociolinguistic and lexical gains. In contrast, evidence of grammatical gains is less forthcoming, being limited to Freed's (1990) findings in the case of her low-proficient learners of French. A possible reason for such discrepancies in findings across different components of the learner's linguistic repertoire in the L2 concerns the fact that strong emphasis is typically placed on the learner's grammatical skills during classroom instruction, such that the classroom may facilitate greater feedback on the learner's grammatical skills. In contrast, other areas of the learner's linguistic repertoire may be more sensitive to more informal contact with the L2, such as



in the TL community. For example, Dewaele (1992, Dewaele and Regan 2002) finds that the level of use of sociolinguistic markers by his instructed Dutch-speaking learners of French varies as a function of their level of informal contact with the L2 outside the classroom, such as in terms of their level of media access and use of the L2 in everyday interaction.

However, whilst it may indeed be true that the foreign language classroom lends itself more properly to more increased development on the learner's grammatical skills, another explanation of the discrepancies in findings may relate to limitations in the type of analysis carried out by previous research on the study abroad learner's grammatical development: in general, the studies reviewed attempt to capture the learner's grammatical development in a very global way. For example, by investigating grammatical development across the learner's grammatical repertoire in the L2, the studies offer a limited picture of how development may vary across the different components which constitute the learner's grammatical competence, such as in the expression of temporality, spatiality and reference, amongst others. As Coleman (1998) points out, it is highly likely that development occurs at different rates on such different components, such that development may be more evident in some areas than in others. He therefore calls for research which investigates specific components of the learner's grammar, rather than investigations of the learner's grammar as a general entity.

Furthermore, by investigating grammatical development as expressed by changes in grammatical accuracy in relation to prescriptive norms, such studies fail to reveal the intrinsic detail behind development on use of specific forms in terms that are interlanguage-specific, and therefore less describable in relation to the TL. In an attempt to identify the specificity of the learner's grammatical development during study abroad as opposed to during foreign language instruction, the study presented here attempts to provide a more fine-grained picture of the complexity of such development. Indeed, that development is often not evident, given the dominant linguistic variation which characterises the learner's use of the L2 grammatical morphology. This study adopts variation theory, which has proved particularly effective in capturing the linguistic variation characteristic of the native speaker's language use, as exemplified by Guy (1990), Labov (1994), and Sankoff (1980), amongst others. Furthermore, variation theory has also illuminated the linguistic system at work in non-standard speakers, such as in the case of regional dialects, pidgins and creoles, and in language contact situations<sup>6</sup>. The L2 learner is also a non-standard speaker, and therefore, variation theory may be potentially advantageous in discriminating between the development which the instructed and study abroad learners



undergo<sup>7</sup>. The study presented here builds on previous preliminary work by applying variation theory to a comparative analysis of grammatical development in a study abroad context and in the foreign language classroom, in an attempt to identify potential changes in the characteristics of a particularly variable aspect of the L2 learner's grammatical skills, namely the expression of past time in TL French.

### **3. Study**

The study is based on a cross-sectional investigation of 18 Irish university learners of French who, at the time of the study, were in their early 20s. The learners' sociobiographic and linguistic characteristics match those which Bartning (1997) identifies as defining the advanced instructed learner, such as prior learning experiences, motivation, and reasons for learning the L2, on the one hand, and level of linguistic development, on the other<sup>8</sup>. For example, in terms of their level of instruction, the learners had previously been learning French for 5-6 years prior to university, where they had completed 2-3 years of their degree programme, for which they were specialising in French and one other subject. Prior to university, the learners had also studied Irish. In terms of their motivation, the learners can be classed as highly motivated insofar as they had chosen French as part of their BA degree programme for professional reasons, and they hoped to use French during their future careers. Thus, whilst they may have differed in terms of their integrative motivation, the learners were probably similar in terms of their instrumental motivation. In terms of their linguistic profiles, the learners can be considered to have passed the creative stage of IL grammaticalisation insofar as the various TL morphological forms had generally emerged in their IL<sup>9</sup>. They were therefore within the stage of adaptive grammaticalisation, whereby their development related, on the one hand, to the refinement of the functional contexts in which the learners use the particular TL forms and, on the other hand, to gaining increased control over use of those forms in real time.

With regard to their programme of instruction in French, the learners' acquisition of French prior to their university studies was based on a communicative approach, as prescribed in the syllabus for modern language teaching in Irish schools. Such an approach attempts to integrate development on the four language skills, as well as aiming to develop the learner's metalinguistic knowledge of formal aspects of the L2, cultural knowledge of the L2, and strategic competence in learning and using the L2. In the

case of their university programme of studies in French, the learners followed courses in French language, literature, and culture, amounting to approximately 200 contact hours annually. Much of the teaching of the learners' courses was done through the medium of French. As part of the language component of the learners' course of study in French, the learners followed courses in both written and spoken French, where explicit focus on the learners' grammatical skills was balanced by use of authentic written and aural materials to stimulate communicative interaction. Thus, in relation to our study of past time marking by the learners, it is important to note that the use of past time morphology had been a regular topic on their language programme, both in terms of the formal presentation of use of such morphology in TL French, as well as in terms of the integration of opportunities for the practical use of such morphology during the communicative activities in which the learners participated.

For the purposes of the data elicitation, the learners participated in individual sociolinguistic interviews with the anglophone researcher, who demonstrated near-native competence French. Whilst the learners knew the researcher to be a staff member of their academic department, they had no prior contact with the researcher in their academic studies. In all cases, the learners volunteered to participate in the study, the details of which they were not aware. Each interview typically lasted one hour, and followed the network of conversational modules proposed by Labov (1984) to best elicit natural spontaneous speech. However, the modules were adapted to suit the interests of the learners, and included such topics as holidays, student life, academic studies, relations between anglophone and francophone speakers, crime, and Labov's 'danger of death' module. Although the study presented here specifically investigates variation on the learners' use of past time morphology in TL French, topics of conversation introduced during the interviews also referred to the present and future time. At the end of each interview, the learners completed a sociobiographic questionnaire, as a means of providing supplementary information concerning their language learning experiences, which was also used to check the validity of some of the information they provided during the interviews. The data were transcribed into standard orthography following the transcription conventions proposed by Blanche-Benveniste and Jeanjean (1987).

Following the data elicitation, the learners were classified into three groups, depending on their level of instruction, and prior residence in the TL community. The learners in group 1 were, at the time of the study, about to participate in a study abroad programme in France following two years of instruction at university. Group 2 consists of similar learners in terms of

their level of instruction. However, at the time of the study, these learners had returned from an academic year in France, where they followed a number of courses through French as part of their academic programme of studies, although they did not receive any formal instruction in French. Finally, group 3 consists of learners who had foregone study abroad to complete the next level of instruction. Unlike the learners in group 2, they had no prior residence in the TL community. Thus, the design of the study allows a comparison with group 1 of the effect of a year abroad, as demonstrated by group 2, relative to classroom instruction, as demonstrated by group 3, on the learners' IL development. As the study is cross-sectional, it was important that the learners in groups 2 and 3 were at the same general level of proficiency as group 1 before they respectively studied abroad and received further instruction. This was done by a comparison of results on a general test of linguistic proficiency completed by the learners after two years of instruction.

The study presented focuses on an aspect of the learners' grammatical skills, namely the expression of past time in TL French. The following section will present an overview of the data analysis carried out.

#### 4. Analysis

As part of an analysis of the learners' variable use of verb morphology for the expression of past time, all verb tokens in past time contexts were extracted from the data including the background contextual detail. Tokens of the verbs *être* and *avoir* were excluded from the analysis due to their overuse<sup>10</sup>. The remaining tokens were firstly coded for morphological form. Three forms dominated, namely the unmarked present form, the *passé composé* (PC), and the *imparfait* (IMP)<sup>11</sup>. Given the variation characterising the learners' use of such markers, the tokens were subsequently coded for a number of linguistic factors which were predicted to constrain the learners' choice of aspectuo-temporal marker. Such factors relate both to the linguistic context in which the verb tokens occurred, as well as to the lexical verb itself. In this paper, we shall restrict our investigation to the factors of grammatical aspect, inherent lexical aspect, and discourse grounding, although other factors were also investigated, such as the effect of verb (ir)regularity, syntactic context, and the presence/absence of temporal adverbials<sup>12</sup>. Whilst such factors have been specifically identified in the literature as constraining the *emergence* of past time morphology in learner IL, in this study we are concerned with their effect on the *variable use* of

such morphology by the advanced learner of French, with a view to considering how each factor might promote over-use of a particular form, and the corresponding under-use of the more appropriate form in context<sup>13</sup>. The coding of the data for such factors allows a detailed insight into potential differences between the aspectuo-temporal system of the study abroad learner and the instructed learner, by detailing how the effect of each factor might differ across the learners as a function of their learning environment. In the following, we shall briefly detail the coding procedure in relation to each factor, beginning with the factor of grammatical aspect.

*Factor 1: Grammatical aspect*

The grammatical aspectual value which the morphological markers were used to express in context, was firstly coded in terms of the primary aspectual values expressed by the PC and the IMP in TL French. In the case of the PC, such values concern the perfect and aorist values, whilst in the case of the IMP, a number of imperfective values were identified relating to its continuous, habitual, and progressive values. Whilst the coding of such values was done by taking account of a number of contextual clues, as outlined below, all ambiguous tokens were excluded from the analysis<sup>14</sup>.

Whilst both values of the PC present the event referred to as an unanalysed whole, differences arise concerning the perspective from which the event is viewed. In the case of the perfect value of the PC, the event is viewed from the temporal perspective of the moment at which the event is recounted, as exemplified in (1).

- (1) *Vous avez déjà entendu*  
 You AUX already hear-PC  
 ‘You have already heard.’

Such a value is made explicit by the presence of certain temporal adverbials such as ‘déjà’ and ‘encore’, which emphasize the experiential quality of this value. In contrast, in the case of the aorist value of the PC in (2), the event recounted is related to another moment in the past which functions as a temporal anchor from which the event is viewed.

- (2) *Quand je suis arrivé en France je suis allé à la mairie*  
 When I AUX arrive-PC in France I AUX go-PC to the townhall  
*pour une carte de séjour*  
 for a residency permit  
 ‘When I arrived in France, I went to the town hall for a residency permit.’

Such a value, which is typical of narrative discourse, where a chronological relation holds between each event recounted, is often made explicit by the presence of temporal adverbials such as 'puis' and 'après'. The presence of temporal adverbials implying perfectivity also provided a clue, such as 'pendant trois semaines'.

In contrast with the PC, the IMP in TL French presents the event in terms of its internal temporal development. Furthermore, use of the IMP implies that the event is coreferential to another event, which functions as a temporal anchor from which the event is viewed. However, the degree of coreference between the event and that reference point can vary, giving rise to a range of values, as noted by Kihlstedt (1998).

The continuous value of the IMP, as exemplified in (3), arises when an event is temporally valid at all moments within the temporal framework explicitly stated or implicitly implied in context.

- (3) *Quand j'étais en France j'habitais dans la cité universitaire*  
When I be-IMP in France I live-IMP in the residence academic  
'When I was in France, I lived in the university dorms'.

Such a value is highly characteristic in TL French when the IMP occurs with stative verbs, since such verbs typically do not impose temporal boundaries. Whilst the temporal validity of the event is restricted to the past thanks to the presence of the temporal framework provided, the validity of the event may extend beyond the boundaries of that framework so as to also hold true in the present. In that case, we coded for an unrestricted continuous value, as exemplified in (4)<sup>15</sup>.

- (4) *Un autre bar qui s'appellait Hideout*  
An other bar which RECP call-IMP Hideout  
'Another bar which was called Hideout'.

Similar to the continuous value of the IMP, its habitual value also applies in the case of events whose temporal validity holds throughout the temporal framework given in context.

- (5) *Pour mon travail j'écrivais des lettres tout le temps*  
For my work I write-IMP some letters all the time  
'For my work I was always writing letters'.

However, unlike the continuous value, the habitual value implies a temporal segmentation which emphasizes the repetitive meaning of the event. Such

a value typically occurs when dynamic verbs are present in context. Unlike the habitual value, a frequentative value was coded in cases where the event did not hold true throughout the duration of the temporal framework given in context, but rather was only valid at certain moments, such that the degree of repetition is much greater than implied by the habitual value<sup>16</sup>.

- (6) *L'après-midi je faisais des traductions de temps en temps*  
 The afternoon I do-IMP some translations from time to time  
 'Sometimes I used to do some translations in the afternoon'.

In order to distinguish between the two values, we relied on a number of criteria relating to the lexical verb itself as well as the presence of adverbial complements of frequency. For example, in the examples provided above, we note that the direct object of the verb used is a nominal plural complement, which of itself implies a series of repetitions. However, the adverbial complements of frequency serve to delimit the frequency of repetition of the event.

In contrast with the previous values, which concern the relation of coreference between an event and a temporal framework which implies duration, we coded for a progressive value in cases where the event was coreferential to a punctual reference time.

- (7) *J'ai éteint le chauffage parce que là aussi ça faisait*  
 I AUX quench-PC the heating because there also it make-IMP  
*de la chaleur*  
 some heat  
 'I switched off the heating because that was also making the place hot'.

In this case, the event is viewed from a more punctual perspective, such that a momentary glimpse of the event is provided, whose validity presumably extends beyond the temporal boundaries of that punctual reference time.<sup>17</sup>

The coding of the data for morphological form and aspectual value allows an investigation of the variation in use of the different markers across the different values, such that certain values may favour use of a specific form. Following the coding of the data for grammatical aspect, the data were secondly coded for the co-occurrence of the aspectuo-temporal markers with different verb types, according to their inherent lexical aspect.

*Factor 2: Inherent lexical aspect*

Inherent lexical aspect concerns the semantic features of the verb which distinguish the underlying meaning of a particular verb from another. Four verb types are identified in the commonly used four-way classification of lexical aspect presented by Vendler (1967), and shown in table 1.

Table 1. Examples of semantic verb types

Verb type	French examples	English examples
Stative	Intéresser, vivre, aimer	Interest, live, like
Activity	Jouer, chanter, marcher	Play, sing, walk
Accomplishment	Passer trois semaines là	Spend three weeks there
Achievement	Décider, atteindre, toucher	Decide, reach, touch

Three semantic features distinguish each verb class, namely punctuality, telicity, and dynamicity. Firstly, punctuality implies that the event referred to is either inherently durative or punctual. Secondly, telicity concerns a transition from one state to another, whereby the event referred to implies a temporal endpoint which brings about the transition. The final defining feature is that of dynamicity which refers to the energy required to carry out the event denoted by the verb. The following chart summarizes the distinguishing characteristics.

Table 2. Semantic features of verb types

Characteristic	Stative	Activity	Accomplishment	Achievement
Punctuality	–	–	–	+
Telicity	–	–	+	+
Dynamicity	–	+	+	+

The coding of the data for the co-occurrence of the different aspectuo-temporal markers with the verb types made it possible to identify to what extent use of a particular marker may be more frequent with certain verb types and less frequent with others. The final factor coded in the data concerned the occurrence of the different markers in different discourse grounds.

*Factor 3: Discourse ground*

Discourse ground concerns the narrative function of the lexical verb within narrative discourse, as defined by Klein and vonStutterheim (1987) in their Quaestio Model. For example, if the event occurs in the foreground, the event enters into a relationship of temporal sequentiality with a previously mentioned event as in (8).

- (8) *Enfin je suis descendue en bas et je suis sortie*  
 Finally I AUX descend-PC below and I AUX go out-PC  
 'Finally I went downstairs and I went out.'

That is to say, the foreground implies temporal progression along the time axis, such that foregrounded events constitute the basic story-line of a narrative. In contrast, the background has a more secondary function in narrative discourse, such that it serves to provide supplementary information surrounding the story in the form of a commentary on or explanation of the events recounted in the foreground.

- (9) *Le soleil qui brûlait tout le monde*  
 The sun which burn-IMP all the world  
 'The sun, which was burning everyone.'

Coding of the data for the occurrence of the different aspectuo-temporal markers across discourse grounds allowed an analysis of the extent to which a particular ground favours use of a specific marker, and in so doing, disfavours use of the alternative markers.

Having detailed the range of factors investigated, it is important to point out that the analysis presented here does not pretend to test the hypotheses which have emerged in the literature surrounding the effect of these factors on the emergence of aspectuo-temporal morphology in learner IL. For example, the factor of grammatical aspect has been incorporated into the 'Aspect before Tense' Hypothesis; the factor of inherent lexical aspect has been developed into the Aspect Hypothesis; and finally the Discourse Hypothesis concerns the factor of discourse ground<sup>18</sup>. Rather, in this paper, we are solely concerned with potential differences and similarities between the groups in terms of the underlying patterns which characterise their use of the aspectuo-temporal markers when the various factors are present in context.



## 5. Results

The results will be presented in terms of the differences and similarities which characterise the groups in relation to their use of verb morphology to express past time. Differences between the groups firstly emerge in relation to their relative use of the various forms, as detailed in table 3, where group 1 refers to the pre-study abroad learners; group 2 refers to the study abroad learners; and group 3 consists of the learners who had not participated in study abroad, but rather had proceeded to the next level of instruction.

Table 3. Formal distribution of the aspectuo-temporal markers

Form	Group 1		Group 2		Group 3	
	n	%	n	%	n	%
Present	321	47	180	14	173	21
PC	275	40	688	52	465	56
IMP	90	13	450	34	190	23

In table 3, a more important effect for study abroad emerges on a number of scores, such that differences between the groups are significant ( $p < .001$ ). Firstly, group 2 makes lesser recourse to the present form than the other groups, such that they mark past time with the past time forms proper to a greater extent. In contrast, the effect of further instruction, as manifested by group 3, is not as important. With regard to the past time forms of the PC and the IMP, differences are also evident: use of the IMP is more frequent in the study abroad learners of group 2. Once again, the effect of further instruction is more restricted since differences between level of use of the IMP by groups 1 and 3 are not as important. However, in the case of the PC, study abroad and further instruction have had a similar effect in increasing the learners' use of this form.

Whilst table 3 simply presents the formal occurrence of the markers in past time contexts, table 4 considers the extent to which discrepancies occur between the learners' formal use of each form, and the functional value expressed in context, by detailing the learners' accuracy of use of each form in context. Such an analysis took account of a number of contextual factors such as the presence of temporal adverbials and the general background context which pointed to the appropriateness of the form used. Since 'appropriateness' is not necessarily either an absolute or wholly

objective phenomenon, the marking as accurate of tokens which did not seem out of place is naturally somewhat subjective. However, ambiguous tokens, whose accuracy of use could not be decided, were not included in the analysis.

*Table 4.* Accuracy of use of the PC and the IMP

Form	Group 1		Group 2		Group 3	
	n	%	n	%	n	%
PC	230	84	613	91	426	83
IMP	70	79	396	92	140	75

Table 4 points to the important discrepancies occurring between the learners' use of each past time form, and the functional value of the context in which the past time forms are produced. That is to say, the learners do not categorically restrict their use of each form to those functional contexts which match the aspectual value expressed by the specific form according to TL prescriptive norms. Rather, the learners under- and over-use each form in contexts where they do not express the aspectual value implied. However, whilst this is true across the groups, the extent to which it is true differs across the groups in the case of each form. A more significant effect for study abroad is noted, insofar as group 2 demonstrates a higher level of accuracy than the instructed groups ( $p=.013$ ). This is true in the case of both the PC and the IMP. In contrast, the effect for further instruction, as demonstrated by differences between groups 1 and 3, is less considerable. Indeed, little difference is apparent between these groups such that further instruction does not appear to have affected the learners' accuracy of use of the past time forms. In contrast, study abroad has had a more important effect, by increasing the learners' level of accuracy on both the PC and the IMP.

Failure to categorically apply the appropriate form in context implies that the learners under-use each form in contexts where they would be prescribed in TL French, and in so doing, over-use another inappropriate form. The expression of past time therefore constitutes an area of considerable variation in their IL. In the remainder of this section, we will be concerned with an analysis of the effect of a number of contextual factors predicted to constrain the learners' variable use of each form. As previously presented, those factors concern the grammatical aspectual context in which the verb to be marked occurs; the inherent lexical aspect of that verb; and the discourse

ground in which the verb occurs. We shall begin with the factor of grammatical aspect, with a view to considering how use of each form varies across different aspectual values. Table 5 presents, in the case of each group, the relative distribution of each form across the various aspectual values. For the sake of convenience, we have grouped the aorist and perfect values together. This is in spite of the fact that they do not necessarily share the same characteristics in the same way as the values of the IMP each share the common invariant of coreference. (For discussion, see Comrie 1975).

By reading down the table, table 5 presents the occurrence of each marker across the various aspectual values. We shall detail the results by considering how use of each form varies depending on the aspectual value to be marked. We shall be particularly concerned with the question of whether the patterns of use of each form across the values might differ across the groups. In the case of use of the present in past time contexts, we have already noted in table 3 that use of this form is most frequent in group 1, whereas its occurrence is less frequent in the other groups, which produce the past time forms to a greater extent. However, table 5 suggests that, across the groups, use of the present none the less varies as a function of whether the context to be marked is perfective or imperfective: across the values which constitute this distinction, we note that use of the present is more frequent in imperfective contexts than in perfective contexts, such that imperfective contexts constitute a more difficult context for past time marking than perfective contexts, as indicated in the following:

*perfective > imperfective*

Within imperfective contexts, we also note that use of the present is most frequent in unrestricted continuous contexts, such that they are most resistant to past time marking. In contrast, use of the present is least frequent in continuous contexts whose validity is restricted to the past.

With regard to use of the PC, we also note, across the groups, that this form occurs in both perfective and imperfective contexts. However, the learners predominantly use it in perfective contexts. None the less, it is also noticeable that use of the PC in perfective contexts is higher in group 2 than in the other groups, such that the marking of perfectivity poses less difficulty to the study abroad learners than the instructed groups. A final point concerns the learners' over-use of the PC in imperfective contexts: we note, across the groups, that this form tends to be over-used to express the progressive, habitual and frequentative values, whereas it tends not to occur in continuous contexts.

Such patterns of use of the present and the PC point to a number of common tendencies across the groups. Further similarities also emerge in relation to use of the IMP<sup>19</sup>. Whilst this form is also over-used in perfective contexts, we note that it is predominantly used to express imperfectivity, as prescribed in TL French. However, we also note that the occurrence of the IMP in imperfective contexts is not as frequent as the occurrence of the PC in perfective contexts, providing further evidence that imperfective marking constitutes greater difficulty than perfective marking, irrespective of the learners' domain of acquisition.

Furthermore, even when it is used, we note that the IMP is not used equally to express the various aspectual values which constitute imperfective aspect, but rather, certain values constitute easier contexts for use of this form than others. In particular, we note that frequency of use of the IMP follows the same common pattern across the groups, as indicated in the following pattern of use:

*continuous > progressive > habitual > frequentative > unrestricted continuous,*

such that use of the IMP is more frequent in continuous contexts than in progressive contexts, which, in turn, are marked with the IMP to a greater extent than habitual contexts. In contrast, frequentative contexts constitute a more difficult context, whilst unrestricted continuous contexts constitute the most difficult context of all. As previously noted, in those progressive, habitual, and frequentative contexts where the learners do not produce the IMP, they principally overextend the PC. In contrast, continuous contexts attract use of the present to a greater extent. In spite of such similar patterns of use of the IMP across the groups, a more important effect nonetheless emerges for study abroad, insofar as use of the IMP is more frequent in group 2 than in the other groups. The effect of further instruction is not as important, since group 3 does not increase its use of the IMP to the same extent as group 2.

In summary, the results concerning the effect of the factor of grammatical aspect on the learners' variable past time marking point to a number of common patterns of use across the groups. In particular, we have noted, on the one hand, that use of present is more frequent in imperfective contexts, and, on the other hand, use of the PC in perfective contexts is more frequent than use of the IMP in imperfective contexts. As such, perfective marking emerges as an easier context than imperfective marking. Furthermore, we have also noted how use of the PC and the IMP varies across the various values which constitute the perfective/imperfective distinction. For example,

Table 5. Relative occurrence of the present, the PC, and the IMP in (im)perfective contexts

Group 1	Present		PC		IMP	
	n	%	n	%	n	%
<i>Perfective</i>						
Perfect	18	28	41	63	6	9
Aorist	75	27	189	69	10	5
<i>Imperfective</i>						
Continuous	19	53	2	5	15	42
Progressive	15	43	7	20	13	37
Habitual	36	49	14	19	24	32
Frequentative	23	43	22	41	9	16
Unrestricted continuous	86	90	–	–	10	10
Group 2	Present		PC		IMP	
	n	%	n	%	n	%
<i>Perfective</i>						
Perfect	4	3	123	97	–	–
Aorist	27	5	484	89	34	6
<i>Imperfective</i>						
Continuous	22	11	–	–	166	88
Progressive	18	28	2	3	44	69
Habitual	24	23	18	18	60	59
Frequentative	34	23	32	22	82	55
Unrestricted continuous	47	47	1	1	53	52
Group 3	Present		PC		IMP	
	n	%	n	%	n	%
<i>Perfective</i>						
Perfect	8	10	71	86	3	4
Aorist	26	7	314	85	32	8
<i>Imperfective</i>						
Continuous	11	13	1	1	75	86
Progressive	5	17	9	33	16	53
Habitual	30	40	13	18	31	42
Frequentative	19	37	19	37	13	26
Unrestricted continuous	63	81	1	1	14	18

*Table 6.* Distribution of aspectuo-temporal morphology with individual verb types

<b>Group 1</b>	Stative		Activity		Accomplishment		Achievement	
	n	%	n	%	n	%	n	%
Present	158	78	65	42	51	38	48	25
PC	23	11	60	38	71	54	121	62
IMP	23	11	31	20	11	8	26	13

<b>Group 2</b>	Stative		Activity		Accomplishment		Achievement	
	n	%	n	%	n	%	n	%
Present	80	24	44	17	27	13	29	6
PC	43	13	109	41	147	69	369	76
IMP	210	63	111	42	39	18	89	18

<b>Group 3</b>	Stative		Activity		Accomplishment		Achievement	
	n	%	n	%	n	%	n	%
Present	74	33	47	29	27	14	22	9
PC	62	28	62	38	145	76	208	80
IMP	88	40	54	33	18	10	30	11

over-use of the PC principally occurs in progressive, habitual and frequentative contexts, whereas it does not occur as frequently in continuous contexts. Finally, with regard to use of the IMP, we have noted a common pattern of use of this form in imperfective contexts, such that continuous and progressive contexts favour marking with the IMP, whereas habitual and frequentative contexts, and unrestricted continuous contexts are more resistant to marking with this form. In spite of such common underlying patterns of past time marking, we have none the less noted a more important effect for study abroad, insofar as the study abroad learners attain a higher level of accuracy in their marking of past time. On the one hand, they make less frequent recourse to the present in past time contexts. On the other hand, they make more frequent use of the PC in perfective contexts, and of the IMP in imperfective contexts than the instructed groups, with corresponding less frequent over-use of the PC in imperfective contexts and of the IMP in perfective

contexts. Such tendencies are reflected in their higher level of accuracy of use of the past time forms. In contrast, the effect for classroom instruction is less important, since group 3 does not demonstrate the same level of development.

Having considered how the factor of grammatical aspect constrains past time marking by the learners, it remains for us to consider the factors of inherent lexical aspect and discourse grounding. We shall begin with the former. The results are presented in table 6, which details the distribution of the aspectuo-temporal forms within each category of verb types.

By reading across the table, one notes how use of each form varies across the verb types, along the lines predicted by the Aspect Hypothesis (Andersen 1991). The effect of inherent lexical aspect on the learners' past time system also suggests a number of similarities across the groups. For example, in the case of the present, we note that occurrence of this form is generally more frequent with stative and activity verbs than with accomplishment and achievement verbs, as presented in the following pattern of use of the present:

*stative > activity > accomplishment > achievement*

In contrast, occurrence of the PC demonstrates the opposite tendency, whereby this form is more frequent with accomplishment and achievement verbs, whereas stative and activity verbs are more resistant to marking with this form, as reflected in the following pattern of use of the PC:

*achievement > accomplishment > activity > stative*

Finally, use of the IMP also demonstrates a similar pattern of use across the groups, such that it is generally favoured with stative and activity verbs, whereas accomplishment and achievement verbs are more resistant to marking with this form, as reflected in the following pattern of use, which is similar to that of the present, but diametrically at odds with that of the PC:

*stative > activity > accomplishment > achievement*

Whilst these tendencies generally hold across the groups, group 1 diverges slightly on use of the IMP: use of this form is slightly less frequent with stative verbs than with activity verbs. However, such divergence can be accounted for on two counts. Firstly, given the limited use of the IMP by this group and corresponding overuse of the present in imperfective contexts,

as noted in table 3, it may be more appropriate to consider use of the present and of the IMP together, in which case it becomes apparent that the effect of verb type is similar to the other groups. Furthermore, in the case of the general category of atelic verbs, namely stative and activity verbs, use of the IMP is greater than with telic verbs, namely accomplishment and achievement verbs, such that the tendency for atelic verbs to occur with the IMP to a greater extent than telic verbs is similar to groups 2 and 3.

Whilst such patterns of use of the aspectuo-temporal markers with the verb types are similar across the groups, differences between the groups simply concern the learners' level of use of each form, rather than the underlying patterns of use of each form across the verb types. The final factor to be examined is that of discourse grounding, results for which are detailed in table 7, which presents the distribution of the markers within each ground.

*Table 7.* Distribution of the aspectuo-temporal markers according to discourse ground

<b>Group 1</b>	Foreground		Background	
	n	%	n	%
Present	34	27	46	43
PC	79	62	43	40
IMP	14	11	18	17

<b>Group 2</b>	Foreground		Background	
	n	%	n	%
Present	15	5	40	15
PC	300	89	95	35
IMP	20	6	133	50

<b>Group 3</b>	Foreground		Background	
	n	%	n	%
Present	13	7	36	21
PC	166	82	79	46
IMP	22	11	58	33



By reading across the table, we note the relative distribution of each form across discourse grounds. Once again, it becomes clear that use of each form is not uniform across discourse grounds, but rather use of each form is favoured in a particular ground, and relatively avoided in the other. For example, in the case of the present, we note that use of this form is more frequent in the background than in the foreground. As such, the foreground attracts marking with the past time forms proper, whereas the background is more resistant to marking with those forms. This is true across the groups, as reflected in the following pattern of use of the present in past time contexts:

*background > foreground*

In the case of use of the past time forms, similar patterns also emerge across the groups. The PC is favoured in the foreground, whereas the IMP is favoured in the background. However, given that the IMP does not express temporal movement in TL French, as implied by foregrounded events, such a tendency is not surprising.

Given the similarities in the effect of discourse ground across the groups, differences between the groups once again simply concern the extent to which they use the forms, rather than their underlying patterns of use of the forms: group 2 shows a higher level of development by using the PC in the FGR and the IMP in the BGR to a higher extent than group 3, which, in turn, uses these forms in these contexts to a higher extent than group 1.

## **6. Discussion and conclusions**

With a view to distinguishing the effect of study abroad and classroom instruction on the advanced learner's grammatical development, this study has adopted variation theory to provide a more detailed insight into the learner's grammar than that provided by previous study abroad research. That research has typically investigated grammatical development simply in terms of changes in the learner's accuracy of use of the various forms which constitute the TL grammar. However, by simply considering development in terms of accuracy, those studies have tended not to take account of the considerable variation which characterizes the learner's use of TL forms, such that underlying development is not always evident. In contrast to that research, this study has investigated a specific aspect of the learner's grammar, with a view to identifying the characteristics of the variation surrounding the learners' marking of past time, in terms of their contextual use of the

aspectuo-temporal forms when certain factors are present/absent in context. That is to say, the study considers the detail of the 'texture' of the learners' grammar in different contexts of acquisition. The results point to a number of similarities and differences characterising the learners' past time system in different contexts of acquisition.

We will firstly discuss the results from the perspective of the differential effect of study abroad as opposed to classroom instruction on the learners' development in the expression of past time. Differences between the groups principally concern their level of use of the aspectuo-temporal markers. In this regard, a more beneficial effect for study abroad emerges, insofar as group 2 marks past time with the past time forms to a greater extent than the instructed groups. That is to say, group 2 makes less recourse to the present form in past time contexts. In contrast, the effect for further instruction is less important, since group 3 has not reduced its use of the present to the same extent as group 2. With regard to use of the past time forms proper, differences between the groups have also become apparent. In particular, we have noted the more important effect of study abroad in increasing the study abroad learners' use of the IMP relative to the instructed learners. However, no differences were noted in use of the PC by groups 2 and 3, such that the effect of study abroad is relatively similar as classroom instruction. Such differences in the study abroad learners' general use of the aspectuo-temporal forms correspond to changes in their accuracy of use of the PC and the IMP in context. They demonstrate a higher level of accuracy of use on both the PC and the IMP. Such increased accuracy is evident across the range of aspectual values expressed by the PC and the IMP in TL French. In contrast, no differences emerge between the instructed learners in groups 1 and 3, such that they do not differ in their general accuracy of use of the past time forms.

In spite of the more important effect outlined for study abroad relative to instruction in terms of the learners' level of use of the aspectuo-temporal markers, as well as in terms of their level of use of these forms in specific aspectual contexts, a number of similarities are also evident across the groups. Such similarities principally concern the learners' contextual use of the aspectuo-temporal forms, when certain factors are present in context. That is to say, a similar effect has been noted across the groups for the various independent factors constraining the learners' use of those forms in context. For example, in the case of the factor of grammatical aspect, we have noted across the groups that use of the present is more frequent in imperfective contexts than in perfective contexts. As such, perfective contexts are more likely to be marked for past time than imperfective contexts. Even

when the learners do produce the past time forms, we have also noted that use of the PC in perfective contexts is more frequent than use of the IMP in imperfective contexts. It follows that imperfective marking constitutes a more difficult entity than perfective marking, irrespective of the learners' level of instruction and learning environment. Moreover, in the case of the various imperfective values, we have also noted similar tendencies across the groups. In particular, given that use of the IMP is more frequent in continuous contexts, which are not valid at the speech time, and in progressive contexts, marking of such values poses less difficulty to the learners than marking of other imperfective values. Such patterns of use of the aspectuo-temporal markers point to a number of similarities between the learners in terms of how they use those markers in context. Thus, in spite of differences in their general level of use of the markers, the level of difficulty posed by the marking of various aspectual values appears to be relatively similar across the groups.

Similar underlying patterns of use also emerge in the case of the other factors investigated. For example, in the case of the factor of inherent lexical aspect, we have also noted that use of each form is not equally distributed across the various verb types. Rather, in the case of the present, use of this form is more frequent with stative and activity verbs than with accomplishment and achievement verbs, such that the former verb types resist past time marking, whereas the latter favour such marking. In particular, when marked for past time, accomplishment and achievement verbs tend to occur with the PC, whereas stative and activity verbs tend to occur with the IMP. Similarly, in the case of the factor of discourse grounding, use of the present is more frequent in background clauses than in foreground clauses. It follows that background are more resistant to past time marking, whereas foregrounded events tend to be marked with the PC.

In summary, the differences and similarities noted between the learners point to the complexity of identifying, in absolute terms, a more beneficial effect for study abroad than classroom instruction on development of the L2 learner's grammatical skills. The similar relative distribution of each form when various contextual factors are present in context points to important similarities in the learners' underlying patterns of past time marking, irrespective of their level of instruction, and their learning environment. Study abroad seems therefore to simply affect the learners' general level of use of the various forms as opposed to the patterns of variation behind their use of these forms in context.

Given that the differential effect of study abroad and classroom instruction simply entails a readjustment in the learners' level of use of the markers,

as opposed to how they use the markers, the results point to the general universality of the effect of the various factors found to constrain the learners' variable use of those markers, irrespective of their level of instruction and learning environment. Such findings are reiterated in investigations of both similar and different aspects of the learner's linguistic repertoire in the L2. For example, Bayley (1994: 176) finds a similar effect for a number of contextual factors on the variable marking of past time by his Chinese advanced learners of English, despite differences in their level of use of past time markers. Citing Labov (1972: 121), he concludes that "acquisition, or movement toward target language norms, consists of adjusting the input probability, or 'particular level of usage'. Such adjustments do not perturb 'the uniformity of the abstract patterns of variation'".

Regan (1996: 261) draws a similar conclusion in her investigation of the effects of study abroad on the variable deletion of the sociolinguistic marker 'ne' by Irish learners of French. She notes that, whilst the learners dramatically increased their rate of deletion of this form, no changes occurred in the effect of the linguistic factors which conditioned the learners' use of this marker before and after a stay in France. She concludes that "the stay in the native speech community makes virtually no difference to certain structural features in the learner language", where structural features, in this context, refer to the underlying patterns behind the learners' use of the marker concerned.

Regan's conclusion that study abroad does not affect her learners' purely structural skills, as opposed to sociolinguistic skills, has generally been interpreted to support the claim that study abroad makes no difference to the learner's grammatical skills<sup>20</sup>. However, such a claim makes no distinction between the learner's structural skills, on the one hand, and the learner's grammatical skills, on the other. Indeed, the results presented here point to the need to make such a distinction, since our results point to a difference in how each skill is affected during study abroad and in the foreign language classroom. The term 'structural skills' seems to be generally used to refer to the patterns of use of the TL forms when various contextual factors are present in context, such as grammatical aspect, inherent lexical aspect, and discourse ground, as investigated in this study. From this point of view, our results corroborate those presented by Regan, by suggesting that study abroad and classroom instruction do not entail radical differences in the effect of those factors on the learners' variable use of the past time markers. In other words, the learner's grammar is not 'restructured' in terms of how the learners use the past time forms in context – the patterns underlying their use of these forms remain relatively similar. However, the learners'

grammatical skills are none the less enhanced by their study abroad sojourn, where 'grammatical skills' is simply used to refer to the learners' ability to actually use the TL grammatical morphology in real time. Grammatical gains are evident, insofar as the learners' level of use of the aspectuo-temporal markers differs as a function of whether they have studied abroad or not. Thus, as acquisition proceeds in terms of changes in the learner's level of use of particular TL forms, the 'underlying structure' of the learner's grammar remains relatively unchanged.

In conclusion, the results point to the complexity of distinguishing between linguistic development in different contexts of acquisition. However, the conflicting results presented here concerning the effects of study abroad on the L2 learner's *structural* skills, on the one hand, as opposed to *grammatical* skills, on the other, point to the important insight that study abroad research can offer into the issue of 'context of acquisition' in SLA research.

### Acknowledgements

This study is based on a project which investigates the acquisition of French by Irish university learners of French. The project was funded by the Association for French Language Studies (AFLS); the Irish Council for Research in the Humanities and Social Sciences; and University College, Dublin, whose support I gratefully acknowledge. I also thank the anonymous reviewers for their comments on an earlier draft of this paper. Any shortcomings naturally remain my own.

### Notes

1. Examples of such programmes include the Erasmus, Lingua, Socrates and Tempus programmes, funded by the European Union, and the Canadian inter-provincial exchanges, American Field Service Programs, and Peace Corps immersion programs in North America.
2. For example, statistics produced by the European Union indicate that the number of learners on non-language-oriented programmes participating in Erasmus exchange programmes alone increased from 2,819 to 66,834 in the period 1987–1998. For an analysis of trends, see deWit (1995).
3. In this regard, statistics produced by the European Union indicate that the number of modern language students participating in the Erasmus exchange programme increased from 631 in 1987/8 to 16,125 in 1997/8.

4. For a comprehensive review of such issues in study abroad research, see Coleman (1997, 1998).
5. For overviews of research on study abroad, see Coleman (1997, 1998), Freed (1995a, 1998), and Howard (2001a, 2001b, 2002c).
6. See, for example, Trudgill (1986), Rickford (1987), and Poplack (1997). For an overview of variation theory, see Poplack (2001).
7. For discussion of variation theory within SLA research, see Howard (2002c), Preston (1995), Regan (1990), Tarone (1997, 2000) and Young (1999). For examples of the application of variation theory to the study of IL variation, see Bayley (1994), Regan (1996), and Young (1996), amongst others.
8. For discussion of the advanced learner variety, see also Howard (1998, 1999).
9. For discussion of IL grammaticalisation, see Giacalone Ramat (1992) and Housen (1997).
10. For a lexical analysis of use of these forms in the L2 learner's past time system, see Howard (2002b).
11. For an analysis of the learners' use of the *plus-que-parfait*, see Howard (2004).
12. For an analysis of these factors, see Howard (2002c).
13. For an overview of the literature on the acquisition of tense-aspect morphology in learner IL, see Bardovi-Harlig (2000), Howard (2002c), and Noyau (2000).
14. For an in-depth presentation of the coding decisions taken, see Howard (2002c).
15. Indeed, whilst it is true that either the present or the IMP may be used in such a case, depending on whether the speaker wishes to assert the validity of the event at the present time or in the past, Kihlstedt (1998) notes that the native speaker tends to use the IMP. For discussion of the distinction, see Klein (1994) and Noyau (2000).
16. For discussion of the habitual and frequentative values, see Bybee et al. (1994) and Kihlstedt (1998).
17. For discussion of the progressive value, see Giacalone Ramat (1995) and Kihlstedt (1998).
18. See, for example, Andersen (1991) and Bardovi-Harlig (1995). For a discussion of the factor of inherent lexical aspect from the point of view of the aspect hypothesis, see Howard (2002a).
19. For an in-depth analysis of the acquisition and use of the IMP by the L2 learner of French, see Howard (to appear).
20. See Coleman (1995) and Freed (1998).

## References

- Andersen, Roger  
 1991 Developmental sequences: The emergence of aspect markers in second language acquisition. In *Crosscurrents in Second Language Acquisition and Linguistic Theory*, Thom Huebner and Charles Ferguson (eds.), 305–324. Amsterdam/Philadelphia: Benjamins.
- Bardovi-Harlig, Kathleen  
 1995 A narrative perspective on the development of the tense/aspect system in second language acquisition. *Studies in Second Language Acquisition* 17: 263–291.  
 2000 *Tense and Aspect in Second Language Acquisition: A Study of Form, Meaning, and Use*. (Language Learning Monograph Series 2.) Oxford: Blackwell.
- Bartning, Inge  
 1997 L'apprenant dit avancé et son acquisition d'une langue étrangère. Tour d'horizon et esquisse d'une caractérisation de la variété avancée. *Acquisition et Interaction en Langue Étrangère* 9: 9–50.
- Bayley, Robert  
 1994 Interlanguage variation and the quantitative paradigm: Past tense marking in Chinese-English. In *Research Methodology in Second Language Acquisition*, Elaine Tarone, Susan Gass and Andrew Cohen (eds.), 157–181. Mahwah, N.J.: Lawrence Erlbaum.
- Blanche-Benveniste, Claire and Colette Jeanjean  
 1987 *Le Français Parlé: Transcription et Edition* [Spoken French: Transcription and Editing]. Paris: Didier Erudition.
- Bryant, De  
 1995 Survival of the interventionist: The personal cost of immersion and social change. *Frontiers: The Interdisciplinary Journal of Study Abroad* 1.
- Bybee, Joan, Revere Perkins and William Pagliuca  
 1994 *The Evolution of Grammar. Tense, Aspect, and Modality in the Languages of the World*. Chicago: University of Chicago Press.
- Coleman, James  
 1995 The current state of knowledge concerning student residence abroad. In *'The Year Abroad': Preparation, Monitoring, Evaluation*, Gabrielle Parker and Annie Rouxville (eds.), 17–42. London: Association for French Language Studies/Centre for Information on Language Teaching.  
 1997 Residence abroad within language study. *Language Teaching* 30: 1–20.  
 1998 Language learning and study abroad: The European perspective. In *Frontiers: The Interdisciplinary Journal of Study Abroad. Special Issue: Language Learning in a Study Abroad Context IV*, Barbara Freed (ed.), 167–203.



DeKeyser, Robert

- 1991 Foreign language development during a semester abroad. In *Foreign Language Acquisition Research and the Classroom*, Barbara Freed (ed.), 104–119. Lexington, M.A.: D.C. Heath & Co.

Dewaele, Jean-Marc

- 1992 L'omission du 'ne' dans deux styles oraux d'interlangue française. *Journal of Applied Linguistics* 7 (1): 3–17.

Dewaele, Jean-Marc and Vera Regan

- 2002 Maîtriser la norme sociolinguistique en interlangue française: le cas de l'omission de 'ne'. *Journal of French Language Studies* 12 (2): 123–148.

deWit, Hans

- 1995 Education and globalization in Europe: Current trends and future developments. *Frontiers: The Interdisciplinary Journal of Study Abroad* I.

Dietrich, Rainer, Wolfgang Klein and Colette Noyau (eds.)

- 1995 *The Acquisition of Temporality in a Second Language*. Amsterdam/Philadelphia: Benjamins.

Freed, Barbara

- 1990 Language learning in a study abroad context: The effects of interactive and non-interactive out-of-class contact on grammatical achievement and oral proficiency. In *Linguistics, Language Teaching and Language Acquisition: The Interdependence of Theory, Practice and Research*, James Atlatis (ed.), 459–477. Washington D.C.: Georgetown University Press
- 1991 Introduction. In: Barbara Freed (ed.), *Foreign Language Acquisition Research and the Classroom*. Lexington, M.A.: D.C. Heath & Co.
- 1995a Language learning and study abroad. In: Barbara Freed (ed.), *Second Language Acquisition in a Study Abroad Context*. 3–33. Amsterdam/Philadelphia: Benjamins.
- 1995b What makes us think that students who study abroad become fluent? In *Second Language Acquisition in a Study Abroad Context*, Barbara Freed (ed.), 123–148. Amsterdam/Philadelphia: Benjamins.
- 1998 An overview of issues and research in language learning in a study abroad setting. In: Barbara Freed (ed.), *Frontiers: The Interdisciplinary Journal of Study Abroad. Special Issue: Language Learning in a Study Abroad Context* IV: 31–60.

Freed, Barbara, Nicole Lazar and Sufumi So

- 1998 Fluency in writing: Are there differences between students who have been abroad and those who have not? Paper presented to the Modern Language Association, San Francisco, December 1998.

Gass, Susan

- 1990 Second and foreign language learning: Same, different or none of the above? In *Foreign Language Learning: A Research Perspective*, Bill



- VanPatten, Trisha Dvorak and James Lee (eds.), 34–44. Cambridge, M.A.: Newbury House.
- Giacalone Ramat, Anna
- 1992 Grammaticalization processes in the area of temporal and modal relations. *Studies in Second Language Acquisition* 14 (3): 279–319.
  - 1995 L'expression de l'aspect progressif en italien langue seconde. *Acquisition et Interaction en Langue Etrangère* 5: 47–78.
- Guntermann, Gail
- 1992 An analysis of interlanguage development over time: Part II, *ser* and *estar*. *Hispania* 75: 1294–1303.
  - 1995 The peace corps experience: Language learning in training and in the field. In *Second Language Acquisition in a Study Abroad Context*, Barbara Freed (ed.), 149–169. Amsterdam/Philadelphia: Benjamins.
- Guy, Greg
- 1980 Variation in the group and individual: The case of final stop deletion. In *Locating Language in Space and Time*, William Labov (ed.), 1–36. New York: Academic Press.
- Hannigan, Terence
- 2001 Work abroad and career development. *Frontiers: The Interdisciplinary Journal of Study Abroad* VII.
- Housen, Alex
- 1997 Le processus de grammaticalisation dans le développement de l'interlangue. *Travaux de Linguistique* 36: 209–222.
  - 1998 Facteurs sémantico-conceptuels et discursivo-fonctionnels dans le développement des systèmes temporo-aspectuels: Aperçu de l'acquisition de l'anglais comme langue étrangère. In *Temps et Discours*, Svetlana Voegleer, Andrée Borillo, Carl Vetters and Marcel Vuillaume (eds.), 257–279. Louvain-la-Neuve: Peeters.
  - 2002 Paramètres liés à l'environnement dans l'apprentissage guidé d'une langue seconde. *Acquisition et Interaction en Langue Etrangère* 16: 23–46.
- Howard, Martin
- 1998 Quasi-bilingual? The advanced learner in second language acquisition research. *PaGes* 5: 63–77.
  - 1999 The advanced learner: Poor relation of the second language acquisition family?. *Cahiers AFLS* 5 (1): 7–26.
  - 2001a The effects of study abroad on the L2 learner's structural skills: Evidence from advanced learners of French. In *EuroSLA Yearbook 1*, Susan Foster-Cohen and Anna Nizegoredzew (eds.), 123–141. Amsterdam/Philadelphia: Benjamins.
  - 2001b Foreign language learners who venture abroad are better than learners who don't, right?: On the role of 'study abroad' in second language acquisition. *PaGes* 7: 75–87.

## Howard, Martin

- 2002a Prototypical and non-prototypical marking in the advanced learner's aspectuo-temporal system. In *EuroSLA Yearbook 2*, Susan Foster-Cohen, Tanya Ruthenberg and Marie-Louise Poschen (eds.), 87–113. Amsterdam/Philadelphia: Benjamins.
- 2002b L'acquisition des temps du passé en français par l'apprenant dit avancé: Une approche lexicale. In *Les Temps du Passé Français et leur Enseignement*, Emmanuelle Labeau and Pierre Larrivée (eds.), 181–204. (Cahiers Chronos 9.) Amsterdam: Rodopi.
- 2002c The effects of study abroad on the variable marking of past time by the advanced learner of French. A comparative analysis of Irish study abroad learners and instructed learners. Ph.D. dissertation, Department of French, University College, Dublin.
- 2004 The emergence and use of the *plus-que-parfait* in advanced French interlanguage. In *Focus on French as a Foreign Language: Multidisciplinary Perspectives*, Jean-Marc Dewaele (ed.), Clevedon: Multilingual Matters.
- forthc. Les contextes prototypiques et marqués de l'emploi de l'imparfait par l'apprenant du français langue étrangère. In *Nouveaux Développements de l'Imparfait*, Emmanuelle Labeau and Pierre Larrivée (eds.), (Cahiers Chronos.) Amsterdam/New York: Rodopi.

## Huebner, Thom

- 1995 The effects of overseas language programs: Report on a case study of an intensive Japanese course. In *Second Language Acquisition in a Study Abroad Context*, Barbara Freed (ed.), 171–193. Amsterdam/Philadelphia: Benjamins.

## Ife, Anne, Gemma Vives Boix and Paul Meara

- 2000 The impact of study abroad on the vocabulary development of different proficiency groups. *Spanish Applied Linguistics* 4 (1): 55–84.

## Kihlstedt, Maria

- 1998 La Référence au Passé dans le Dialogue. Etude de l'Acquisition de la Temporalité chez des Apprenants dits Avancés de Français. PhD Thesis, Stockholm University/Akademitryk.

## Klein, Wolfgang

- 1994 *Time in Language*. London: Routledge.

## Klein, Wolfgang and Christine von Stutterheim

- 1987 Quaestio und referentielle Bewegung in Erzählungen. *Linguistische Berichte* 109: 163–183.

## Labov, William

- 1994 *Principles of Linguistic Change. Volume 1: Internal Factors*. Oxford: Blackwell.

## Lafford, Barbara

- 1995 Getting into, through, and out of a situation: A comparison of communicative strategies used by students studying Spanish abroad and 'at

- home'. In *Second Language Acquisition in a Study Abroad Context*. Barbara Freed (ed.), 97–121. Amsterdam/Philadelphia: Benjamins.
- Marriot, Helen
- 1995 The acquisition of politeness patterns by exchange students in Japan. In *Second Language Acquisition in a Study Abroad Context*, Barbara Freed (ed.), 197–224. Amsterdam/Philadelphia: Benjamins.
- Möhle, Dorothea and Manfred Raupach
- 1983 *Planen in der Fremdsprache*. Frankfurt: Peter Lang.
- Mourelatos, Alexander
- 1981 Events, processes, states. In *Tense and Aspect*, P. Tedeschi and A. Zaenen (eds.), 191–212. (Syntax and Semantics 14.) New York: Academic Press.
- Noyau, Colette
- 2000 Les relations temporelles dans la dynamique des lectures d'apprenants. Working paper prepared for the 3<sup>rd</sup> Euroconference 'The Dynamics of Learner Varieties', San Feliu (Spain), October 2000.
- Poplack, Shana
- 1997 The sociolinguistic dynamics of apparent convergence. In *Towards a Social Science of Language. Volume II: Social Interaction and Discourse Structures*, Greg Guy, Crawford Feagin, Deborah Shiffrin and John Baugh (eds.), 285–309. Amsterdam/Philadelphia: Benjamins.
- 2001 Variability, frequency and productivity in the irrealis domain of French. In: Joan Bybee and Paul Hopper (eds.), *Frequency Effects and Emergent Grammar* Amsterdam/Philadelphia: Benjamins.
- Preston, Dennis
- 1995 Variationist perspectives on second language acquisition. In *Second Language Acquisition and Linguistic Variation*, Robert Bayley and Dennis Preston (eds.), 1–45 Amsterdam/Philadelphia: Benjamins.
- Raffaldini, T.
- 1987 Attrition of communicative ability among former year abroad students of French. Ph.D. dissertation, Indiana University.
- Regan, Vera
- 1990 Sociolinguistics and second language acquisition: A variationist perspective. *Teanga: Journal of the Irish Association for Applied Linguistics* 10: 14–24.
- 1995 The acquisition of sociolinguistic speech norms: Effects of a year abroad on second language learners of France. In *Second Language Acquisition in a Study Abroad Context*, Barbara Freed (ed.), 245–267. Amsterdam/Philadelphia: Benjamins.
- 1996 Variation in French interlanguage: A longitudinal study of sociolinguistic competence. In *Second Language Acquisition and Linguistic Variation*, Robert Bayley and Dennis Preston (eds.), 177–201. Amsterdam/Philadelphia: Benjamins.

Regan, Vera

- 1998 Sociolinguistics and language learning in a study abroad context. In *Frontiers: The Interdisciplinary Journal of Study Abroad. Special Issue: Language Learning in a Study Abroad Context IV*, Barbara Freed (ed.), 61–90.

Rickford, John

- 1987 Are black and white vernaculars diverging? *American Speech* 62,1: 55–62.

Ryan, John and Barbara Lafford

- 1992 Acquisition of lexical meaning in a study abroad environment: *Ser* and *estar* and the Granada experience. *Hispania* 75: 714–722.

Salaberry, Rafael

- 1998 The development of aspectual distinctions in academic L2 French. *Canadian Modern Language Review* 54: 508–542.

Sankoff, Gillian

- 1980 *The Social Life of Language*. Philadelphia: University of Pennsylvania Press.

Stephenson, Skye

- 1999 ‘Two cultures under one roof’: Study abroad as a transformational experience and its impact upon exchange students and host nationals in Santiago, Chile. *Frontiers: The Interdisciplinary Journal of Study Abroad* V.

Tarone, Elaine

- 1997 Analysing IL in natural settings: A sociolinguistic perspective on second language acquisition. *Communication and Cognition* 30: 137–149.
- 2000 Still wrestling with ‘context’ in interlanguage theory. *Annual Review of Applied Linguistics* 20: 182–198.

Trudgill, Peter

- 1986 *Dialects in Contact*. Oxford: Blackwell.

VanPatten, Bill

- 1987 Classroom learners’ acquisition of *ser* and *estar*: Accounting for developmental patterns. In *Foreign Language Learning: A Research Perspective*, Bill VanPatten, Trisha Dvorak and James Lee (eds.), 19–32. Cambridge, M.A.: Newbury House.

Vendler, Zeno

- 1967 Verbs and times. In: Zeno Vendler (ed.), *Linguistics in Philosophy*. 97–121. Ithaca, N.Y.: Cornell University Press.

Young, Richard

- 1993 Functional constraints on variation in interlanguage morphology. *Applied Linguistics* 14 (1): 76–97.
- 1999 Sociolinguistic approaches to SLA. *Annual Review of Applied Linguistics* 19: 105–32.

# **The effect of type of acquisition context on perception and self-reported use of swearwords in L2, L3, L4 and L5**

*Jean-Marc Dewaele*

This paper deals with the impact of the acquisition context (instructed, naturalistic, mixed) on two variables reflecting sociolinguistic, sociopragmatic and sociocultural aspects of communicative competence, namely perception of emotional force of swearwords and self-reported language choice for swearing. Data for L2, L3, L4, and L5 were collected from a total of 1039 multilinguals through a web questionnaire containing closed-ended Likert-type questions and open-ended questions. Multivariate analyses revealed that context of acquisition and type of contact (instructed, mixed, naturalistic) had a significant effect on the self-reported use and perceived emotional force of swearwords. The effect was generally stronger for self-reported language choice for use of swearwords than for perception of their emotional force. Multiple regression analyses showed that other variables, namely frequency of use of the language and, to a lesser degree, age of onset are good predictors of perceived emotional force of swearwords and preferred language for swearing.

## **1. Introduction**

Children growing up in a specific social environment develop not only linguistic skills but also the crucial communicative competence, i.e. the sociolinguistic and pragmatic rules of their speech communities. They learn to produce “appropriate” speech in a wide variety of situations with different interlocutors. They realize that words, expressions, utterances have not only propositional meaning or locutionary force, i.e. a certain “face value” (Austin 1962) but also an illocutionary force, referring to the speaker’s intended meaning or “goal” (Leech 1983: 13). The process, as parents and teachers can testify, may take some time but leads ultimately to full competence. No such success is guaranteed for the older instructed second language learner. Grammatical skills can be learned in the classroom, but it is

not the ideal place to acquire sociopragmatic and intercultural competence (Byram 1997, Sercu 2000). The range of registers is restricted, the situation is relatively artificial and the number of potential (native) interlocutors is extremely limited (Mougeon et al. 2002). As a result, classroom learners have incomplete semantic and conceptual representations of target language (TL) words or expressions and would typically be unsure about their precise emotional force and illocutionary effects. They could also transfer discourse patterns and sociopragmatic rules from their L1, which might lead to violations of the TL norms (Kecskes 1998). Anecdotes of this kind of sociopragmatic mishaps are eagerly exchanged among multilinguals. One that stands out in my mind is an episode that happened during a schooltrip to Paris. Our French-Foreign Language teacher had taught us, native speakers of Dutch, a list with colloquial words and expressions in order to “survive” in a linguistic environment reputed for its use of vernacular speech or “argot” (words like *le fric*, *les flics*, *le mec*... ‘money, cops, guy’ that are commonly used across social classes). The trouble began when we sat down on the well-tended lawn in front of the Pantheon, pretending not to understand the sign *Pelouse interdite* (‘Keep off the grass’). An irate guard in his sixties yelled and chased us off the lawn, pointing at the sign. A friend perceived this as a unique opportunity to try out a new speech act in French and uttered the words *Avec ta sale gueule* (literally ‘with your dirty mug’, meaning ‘shut your gob’). He had clearly underestimated the illocutionary effects of the expression. The guard turned red and his medals started clinking. He threatened to call the police and called out to our teacher: *Moi, monsieur, ancien combattant de la dernière guerre, il m’a dit, vous vous rendez compte monsieur, “avec ta sale gueule”! C’est l’honneur de la France qu’il a bafoué monsieur!* (‘He told me, sir, a veteran of the last war, can you imagine sir, ‘with your dirty mug’. It’s the honor of France that he has treated with contempt sir!’). My teacher managed to calm the man down: *C’est un Flamand monsieur, son français est élémentaire, ce qu’il vous a dit ne signifie pas plus pour lui que “moules et frites”.* (‘He is a Fleming sir, what he said to you doesn’t mean much more than ‘mussels and French fries’). The anecdote illustrates the common assumption that for L2 learners “communicative competence may be harder to acquire than linguistic competence” (Davies 2003: 23). Anecdotes might nicely illustrate the phenomenon under consideration, but to draw any firm conclusion one needs a bird’s-eye perspective, i.e. a large corpus, with the proper descriptive and inferential statistics. In the present study, I will use the database on bilingualism and emotions (Dewaele and Pavlenko 2001) to focus on the issue of communicative competence in second or foreign lan-

guages, and more specifically on the perception and self-reported use of swearwords in L2, L3, L4 and L5. I will begin with a discussion of the basic concepts related to communicative competence and the different research designs used in studies on type of contact or acquisition context. I will then survey research that focused on the development of sociolinguistic, sociopragmatic and sociocultural competence after which the hypotheses and the rationale of the present study will be introduced. The methodology of the study is presented in the following section. Next I will present the quantitative analysis of the dependent variables and link them to qualitative data obtained from the same informants. Finally, I will discuss the findings and present some pedagogical implications of the study.

## **2. Some basic concepts in research on types of contact/contexts of acquisition**

The notion of communicative competence was originally proposed by Hymes (1970). The rationale for this notion was the redressing of what Hymes regarded as the narrowness and inadequacy of Chomsky's definition of linguistic competence (Davies 2003: 98). Language users need to be able not only to create and understand grammatical utterances, but also need knowledge about cultural norms in order to judge the social situation correctly so as to produce appropriate speech.

"The position taken up by communicative competence is that knowing what to say is never enough; it is also necessary to know how to say it. And by 'how' is not meant the performing of the speech that is getting the words out; rather what is meant is using the appropriate register, variety, code, script, formula, tone and formality" (Davies 2003: 23).

Hymes' model of communicative competence was further developed for application to foreign language learners (cf. Harley, Allen, Cummins and Swain 1990; Davies 2003). The study of communicative competence encompasses sociolinguistics, sociopragmatics and sociocultural theory. The frontiers between these three disciplines are rather fuzzy. Variationist sociolinguists typically focus on variation patterns, quantifying the frequency of use of a particular linguistic variant and identifying both linguistic sources of variation and extralinguistic variables (gender, social class, group identity, situation, register); sociopragmaticists consider the sociological interface of pragmatics concerned with "the social perceptions underlying participants' interpretation and performance of communicative action" (Kasper and Rose 2001: 2); while researchers using a sociocultural approach investigate



language users' ability "to identify, categorize, perceive and engage in verbal and non-verbal behaviors similarly to other members of a particular speech community" (Dewaele and Pavlenko 2002: 268). Many studies combine at least two of these approaches, defying any clear categorization. The overlap between sociolinguistics, sociopragmatics and sociocultural theory is best illustrated with the previously mentioned incident at the Pantheon. My friend managed to pack an impressive number of sociolinguistic, sociopragmatic and sociocultural errors using a single expression that had no apparent emotional resonance for him. Addressing someone older and of superior status, he should have used the formal form of address "vous" and avoided the colloquial word "gueule"; the expression uttered with the right intonation and pronunciation (which might have convinced the guard that he was dealing with a native speaker) would have been fine to challenge another boisterous schoolboy but it was clearly inappropriate with an adult, particularly an adult who was wearing an uniform, guarding the national shrine for France's great men and who perceived himself as the personification of the republican values of the nation. In other words, that colloquial expression was the wrong thing to say to that particular interlocutor in that particular place.

Lyster (to appear) notes that the teaching of second or foreign languages has a well-known history of explicit rule-based instruction and metalinguistic analysis. This analytic approach to second and foreign language teaching has appeared at odds with basic tenets of communicative language teaching that proscribe explicit instruction in favour of implicit and incidental language learning. Lyster (1994, 1996, 2002) investigated the effectiveness of different types of instruction in classroom contexts in order to demonstrate the usefulness of varying degrees of explicitness and metalinguistic awareness. Lyster (to appear) points out that:

"Questions remain as to how form-focused instruction can best be integrated into communicative contexts and whether explicit knowledge and metalinguistic awareness can, over time and through meaningful practice, become part of a learner's underlying system of implicit knowledge and thus available for spontaneous language production."

Several SLA researchers have reflected on ways to boost language learners' sociolinguistic competence in an instructed context (Cuq 1994; Critchley, 1994; Offord 1994). Valdman (2003) for example has argued in favor of a pedagogical norm that would include not just the standard norm but the multiple norms that coexist in the target language. That would allow learners to shift their norm orientation depending on the situational context and their communicative intent. Valdman thus advocates an interventionist



approach and rejects the laissez-faire attitude of misinterpreted communicatively oriented instruction.

It is clear that materials used in the classroom should reflect real life interactions. Yet, Myers-Scotton and Bernstein (1988) offered proof of the gap between real life and classroom discourse in an analysis in which a textbook dialogue between a man asking for directions and his interlocutor is considerably different from an identical interactional event in the street.

### **3. Previous research on type of contact, context of acquisition and the development of communicative competence in multiple languages**

A majority of studies that considered the effect of context of acquisition have concentrated on fluency, grammatical accuracy, and complexity (cf. Coleman 1996; Ellis 1989, 1990, 1999; Freed 1995; Giles and Coupland 1991; Pica 1983; Towell et al. 1996; Towell and Dewaele, in press). The general finding is that learners who have spent a period abroad have become more fluent, and grammatically more native-like.

Studies of the acquisition of sociolinguistic competence by L2 learners started appearing in the late 1980s and bloomed in the 1990s (Adamson 1988; Adamson and Regan 1991, Bayley 1991; Beebe 1988; Dewaele 1999, Ellis 1999; Regan 1995, 1996; Sankoff et al. 1997; Rehner and Mougeon 1999; Tarone 1988, 1990, 1998; Young 1991, 1999). Here again the general findings were that students who had increased contact with NS of the TL language developed a better understanding of the sociostylistic value of variants and they were better able to use them appropriately in a wider range of situations. Overall however their frequency of use of different variants remained distinct from that of NS control groups. Some of this research focused on the sociolinguistic competence of Anglophone Canadian students in French immersion classes (Blondeau et al. 2002; Harley et al. 1990; Lyster 1994; Lyster and Rebuffot 2002; Tarone and Swain 1995). One consistent finding in these studies is that L2 users from immersion classrooms tend to overuse formal variants and use English for vernacular interactions among themselves:

“In these classrooms, the L2 becomes the superordinate language variety, used predominantly for academic topics in conversations with teachers or for public discourse addressed to the class as a whole. Pre-adolescent socializing will be done in the NL vernacular. This will come about because the learners do not have the L2 vernacular input they need from the L2 interlocutors they need: peers” (Tarone 1998: 433–434).

The year abroad and increased authentic interaction in the TL was shown to have a powerful effect on the development of sociolinguistic competence. Thomas (2002) compared the use of variable pronunciation rules in a group of advanced Canadian English-speaking learners of French who had spent a year in France and a control group who had studied French in their home environment. The study abroad learners had come much closer to native French phonetic norms for liaison and schwa deletion. A year in a French-speaking environment showed similar effects in relation to deletion of the preverbal negative particle *ne* (very common in informal native French) in a longitudinal corpus of French IL of Hiberno-Irish English learners (Regan 1995, 1996, 2002). The studies showed that the learners deleted considerably more in a sociolinguistic interview with the researcher after their stay abroad. In other words, their deletion rates approximated roughly to the native speaker norm. It seems thus “that living abroad for an extended period does something to the learners’ usage which classroom input does not” (Regan 2004: 200). Varbrul analyses showed that while the rate of deletion more than doubled, most of the linguistic factors which condition this deletion remained the same. Regan (2004: 200) therefore concludes that “for these advanced learners of French, their structures in relation to negation remained basically the same, but their sociolinguistic knowledge increased significantly”.

Research into the use of colloquial words in the L2 combines sociolinguistic and sociopragmatic enquiry. Dewaele and Regan (2001) analyzed the proportion of colloquial words (including swearwords) in a cross-sectional corpus of advanced oral French IL of Dutch L1 speakers, in a longitudinal corpus of Hiberno-Irish English L1 speakers, and compared these results with data drawn from a corpus of native speakers of French. Colloquial vocabulary was identified using the stylistic indications of the monolingual French dictionary *Le Petit Robert*. It was hypothesised that authentic interactions in the TL, as well as total immersion in the TL culture, and longer and more intense formal instruction in the TL would be linked to a more frequent use of colloquial vocabulary. The latter factor was found to have no predictive value on the use of colloquial vocabulary in advanced French IL. Only active authentic communication in the TL, especially during the study abroad was found to be linked to an increased use of colloquial vocabulary. It was hypothesised that speakers of intermediate proficiency simply did not know the colloquial words, or lacked the necessary morpho-phonological information at the lexical level. It was also argued that incomplete semantic representation of the words could prevent the production of colloquial words in more advanced speakers, and even highly fluent speak-

ers might either lack information at the conceptual level, or the semantic representation might not be linked to the TL concept, or the scripts where these words could appear might be incomplete or lacking. The authors argued that only prolonged authentic contact with the TL community might allow learners to develop the kind of implicit, proceduralised sociopragmatic knowledge that is stored outside the declarative memory (see Dewaele and Wourm 2002 for an in-depth discussion). The study by Toya and Kodis (1996) on the use of swearwords and the pragmatic use of rudeness in an L2 also found that their dependent variables were clearly linked to the variety of registers in the input and the confidence of the L2 users. They focused on the use of rude expressions as a result of anger among native speakers (NS) of English and NS of Japanese with advanced English proficiency (the latter group providing data for L1 and L2). Participants were presented with five situations in which anger was expected and were asked (1) how they would feel in each situation, (2) how they would or would not express their emotions verbally and/or nonverbally, and finally (3) why they would or would not express themselves in those ways. NS were found to be more expressive although the difference in reactions was smaller than expected. The authors suggest that the lower degree of expressiveness in the L2 could be linked to the more restricted input to which the learners had been exposed (there is little display of anger in the foreign language classroom) and the fact that learners have little confidence in using angry words.

Hoffman Hicks (2000) tested the sociopragmatic skills in French L2 of American students studying abroad over a period of sixteen months. The analysis revealed that the learners did exhibit sociopragmatic development over time but that this development was often slight and limited in scope compared to a first control group of native speakers of French. The findings became more significant, however, when compared with the results of a second, non-native control group, who continued their study in the US (for an in-depth discussion about the development of pragmatic competence in the L2, see Kasper 2001; Kasper and Rose 1999, 2001, 2002).

The sociocultural approach in studies on context of acquisition is best exemplified by the study of Pavlenko (1999). She showed that fluent L2 learners who have only been exposed to classroom TL instruction and L2 users who had lived in the L2 context had different conceptual representations for certain culture-specific notions. Pavlenko argued that conceptual differences between Russian and American English led monolingual Russians and Americans to describe the same two films in very different terms, with Americans privileging the notions of privacy and personal space, which are not part of Russian discourse. These concepts and related

scripts were also found in narratives of Russian L2 users of English who spent between 4 and 7 years in the US. They did not show up however in narratives of advanced L2 learners who had enjoyed up to 10 years of formal instruction but never visited an English-speaking country or had any long-term contact with native speakers of English. These findings suggest that some aspects of sociopragmatic and sociocultural competence may only develop during the long process of second language socialization. A similar point is made in Kecskes and Papp (2000).

It would be an oversimplification however to claim that awareness of sociolinguistic, sociopragmatic and sociocultural rules in the TL cannot be raised in a classroom context. Lyster (1994, 1996) showed that some teaching techniques do accelerate the development of sociolinguistic and pragmatic competence in a TL, albeit not to native speaker levels (see also Ellis 1992; House 1996; Kasper and Rose 2001). A first and crucial step would be to use more authentic material in the classroom, something learners themselves prefer (cf. Chavez, 1998).

The studies reviewed so far have determined competence through production data. Another possible approach is through self-report. Here again researchers have a wide range of methodological approaches at their disposal. Performance data can be matched with self-reports. For instance, Lapkin et al. (1995) considered a wide range of variables and showed that French L2 learners who had stayed for three months with French-speaking families in Quebec reported that their sociolinguistic competence had increased considerably, specifically their ability to express themselves using a vernacular style. Data can be collected solely through interviews with students after their period abroad as in Evans (1988). This study highlighted important differences in self-reported social behaviour of British language students while abroad. While some students used every opportunity to engage in conversations with TL speakers, others avoided contact outside their own linguistic community. This might account for the wide inter-individual variation in the amount of linguistic progress after such a stay abroad (cf. Regan 2002). Testimonies reveal that the period abroad was considered to be crucial for the linguistic development: "If you want to get to the heart of the language, you go out there" (Evans 1988: 43). Students had clearly also developed their sociopragmatic competence in the TL. One student related the following: "The Italians are so different, and if they want something they will go out and get it. I've been taught that you ask for it politely. You realize that unless you do what they do, shout, nothing will come out of it" (Evans 1988: 45). The period abroad seemed to have boosted the students' overall self-confidence, and hence also the their lin-

guistic self-confidence, including the use of shouting! Students can also be asked to fill out questionnaires with open-ended and closed-ended questions relating to different aspects self-perceived performance and competence in the TL language. Such self-report measures have been proved to correlate highly with performance measures of proficiency (Dufour and Kroll 1995; Kroll et al. 2002). They have the added advantage of being easy to collect, enabling researchers to consider larger sample sizes than in research based on production data. Of particular interest for the present study is the research aimed at measuring the perception of emotional intensity of words and expressions in L1 and L2 through a wide variety of methodologies. Bilinguals generally report higher levels of emotional resonance in their dominant language (Gonzalez-Reigosa 1976, Harris, Ayçiçeği and Gleason 2003). Second language learners find it particularly difficult to judge the degree of emotional intensity of speech in the L2 accurately (Rintell 1984, Graham, Hamblin and Feldstein 2001). One possible interpretation is that higher scores on perceived emotional force in an L2 reflects higher levels of sociopragmatic and sociocultural competence (Dewaele to appear a, b). To sum up, the existing body of research on the acquisition of sociolinguistic, sociopragmatic and sociocultural competence seems to suggest that mixed contact (instructional + naturalistic) leads to higher levels of competence compared to purely classroom-based instruction. However, it is difficult to know whether this is also true when intensity and amount of exposure are held constant.

#### **4. Hypotheses**

This study will test the following two hypotheses:

- 1) Mixed and naturalistic learning result in higher levels of perception of emotional force and more frequent use of the TL for swearing than instructed learning.
- 2) A lower age of onset of TL learning and a higher frequency of use of the TL will be linked to higher levels of perception of emotional force and more frequent use (self-reported) of the TL for swearing.

## 5. Rationale for the present study

While one does not doubt the superiority of the mixed acquisition context, there are several methodological and theoretical issues that have not been sufficiently explored or remain unclear with regard to the context of acquisition.

Firstly, analyzing the use of sociolinguistic variants or any variable reflecting sociopragmatic or sociocultural competence in a limited speech sample with a single interlocutor in a relatively formal context might allow a researcher to gather solid empirical information but it allows only a glimpse into the speaker's actual competence. This might not be a major problem when the object of investigation is grammatical competence, but it could be more problematic when the research focuses on variables related to the social aspects of communicative competence. These variables are by definition more dependent on the situation in which the interaction takes place. The use of self-reported data in this study may provide a better reflection of sociolinguistic, sociopragmatic or sociocultural competence as the participant has to condense a life-long communicative history to a single score on the dimension(s) under investigation. Questionnaires used in personality psychology also rely on self-report of usual behavior to determine an individual's position on different personality dimensions.

Secondly, the use of an on-line web questionnaire allowed me to gather data from a very large sample of learners and long-time users of multiple languages (rather than only the L2 as in previous research) from across the world and from a wide age range, i.e. not only from the 18–22 year-olds which have been predominantly used in previous empirical research in applied linguistics and psychology.

I am aware, however, that the present approach is not without its own methodological limitations (cf. Pavlenko 2002). Questionnaires are by nature incomplete as one is forced to find a fine balance between the amount of topics covered and the amount of detail requested while keeping total length under control. Dörnyei (2003: 132) suggests a maximum length of about 4 pages and 30 minutes to complete. Questionnaires with Likert scales responses have been tried and tested extensively in sociopsychological research (cf. Dörnyei 2003). They can provide excellent baseline data, provided they are backed up by different types of data. In the present study, this "different type of data" comes from open-ended questions inquiring into emotion and communicative behaviour.

## **6. Method**

### 6.1. Participants

A total of 1039 multilinguals contributed to the database (731 females, 308 males). The participants spoke a total of 75 different L1s. English speakers represent the largest group:  $n = 303$ ; followed by Spanish:  $n = 123$ ; French:  $n = 101$ ; German:  $n = 97$ ; Dutch:  $n = 76$ ; Italian:  $n = 52$ ; Catalan:  $n = 32$ ; Russian:  $n = 29$ ; Finnish  $n = 28$ ; Portuguese:  $n = 20$ ; Greek:  $n = 15$ ; Swedish:  $n = 15$ ; Japanese:  $n = 11$ ; Welsh:  $n = 10$ . The sample could be described as highly polyglot with 144 self-reported bilinguals, 269 trilinguals, 289 quadrilinguals and 337 pentalinguals. They are also generally highly educated with 115 having a high school diploma, 273 with a Bachelors degree, 308 with a Masters, and 338 with a PhD. A majority of participants ( $n = 837$ ) have a language-related profession. Age ranged from 16 to 70 (Mean: 35.6; SD: 11.3).

### 6.2. Independent variables

Data were gathered through an on-line web questionnaire with 34 questions related to bilingualism and emotion (Dewaele and Pavlenko 2001). The following sociobiographical information was collected: sex, age, education level, ethnic group, occupation, languages known, dominant language(s), chronological order of language acquisition, context of acquisition/type of contact, age of onset, frequency of use and typical interlocutors. L1, L2, L3, L4, L5 was defined in terms of timing and sequence of acquisition rather than in terms of their relative dominance or preference. Thus, the L2 is the second language acquired / learned by the individual, the L3 is the third language etc. A closer look at the ages of onset for learning the L2, L3 revealed that 157 L2 users are in fact “bilingual first language” users, having learned the L2 before the age of 12 months. This represents 15% of the L2 group. Similarly, 19 L3 users are “trilingual first language” users (representing 1.8% of the L3 group). There are no “quadrilingual first language” users. Following Cook’s (2002: 4) recommendation, the term of L2, L3, L4, L5 users is used for the participants, rather than learners.

Although Cook uses “L2” to indicate any language that was learnt after the first one, I prefer to specify whether it concerns L2, L3, L4 or L5 of the participants to avoid confusion when comparing the different groups.

The focus of the present study is on “context of acquisition” or “type of contact with the TL”. This context is of course infinitely richer than the



crude three-fold distinction made in this study between “naturalistic context” (i.e. no classroom contact, only naturalistic communication outside school), “mixed context” (i.e. classroom contact + naturalistic contact) and “instructional context” (i.e. classroom contact only). For instance, as far as “instructional context” is concerned, no further distinction was made between “language-classrooms”, where the target language is the instructional target, and “immersion classrooms”, where the target language primarily serves as the medium for teaching non-language subject matter. Learning and teaching practices at school have considerably evolved over the years, and still vary geographically and socially, but they all share one aspect, namely that learning happens within the confines of classroom walls, in the presence of a teacher and classmates. Similarly, also the notion of “naturalistic context” as used here is a cover term for a wide range of ways in which one can learn a language naturalistically. However, they all have in common that the learning process was not intentionally guided by a particular teacher or program, but developed gradually, spontaneously through interaction with speakers of the TL.

I am not claiming, either, that the independent variables in the study are the only ones to determine the dependent variables. Important additional variables include amount, intensity and duration of TL exposure/contact. The questionnaire used in the present study only enquired about past and present TL use in broad terms (i.e. How frequently do you use each of the languages? Possible answers included: Never, every year, every month, every week, every day, several hours a day. It is hoped that statistical techniques will allow to estimate the proportion of variance in the data that can be attributed to the independent variables under consideration.

### 6.3. Dependent variables

Swearwords are multifunctional, pragmatic units which assume, in addition to the expression of emotional attitudes, various discourse functions. They contribute, for instance, to the coordination of the interlocutors, the organisation of the interaction, and the structuring of verbal exchange; in that, they are similar to discourse markers. They are also used as a linguistic device to affirm in-group membership and establish boundaries and social norms for language use (Drescher 2000). I pointed out in Dewaele (2004b) that there is an interesting paradox concerning the use and perception of swearwords in the L2. They are often among the first ones to be learned in an L2, typically over a drink with NS. Yet, these words rarely in textbooks or



in the classroom discourse because of their offensive character. Moreover, NS often display an ambiguous and proprietary attitude toward NNS using swearwords in “their” language. By doing so, the NNS seem to be claiming in-group membership despite the fact that multiple linguistic indices prove the contrary. A swearword uttered by a NNS will have a different illocutionary effect compared to the same word in the mouth of a NS.

The present study focuses on the self-reported use of swearwords and the perception of their emotional force. The first question was formulated as follows: If you swear in general, what language do you typically swear in? Possible answers on a 5-point Likert scales included: never=1, rarely=2, sometimes=3, frequently=4, all the time=5. Information was collected for L1, L2, L3, L4, and L5. The second question was: Do swear and taboo words in your different languages have the same emotional weight for you? Participants were again asked to circle the appropriate number: 1=does not feel strong, 2=rather weak, 3=fairly strong, 4=strong, 5=very strong.

#### 6.4. Research design

The main independent variable is the type of contact or acquisition context of L2, L3, L4 and L5 (instructed, mixed or naturalistic). Multivariate analyses of variance (MANOVA) and Scheffé post-hoc tests were used to check for intergroup differences. Sample sizes may vary across the analyses because some participants did not provide data for all the dependent variables. The questions related to swearing were left blank more often because, as one participant put it “I don’t swear”.

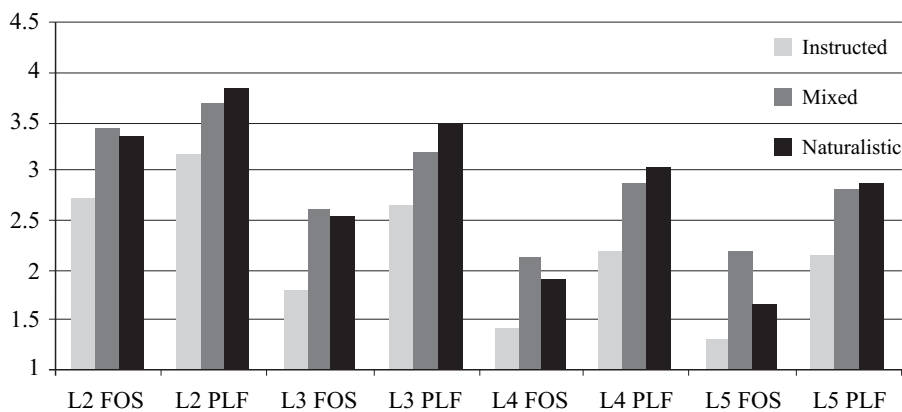
Pearson correlation analyses were used to check whether the dependent variables were linked to other independent variables such as age of onset of learning and frequency of communication in the language. Standard multiple linear regression was used to examine the hypothesized relationships between a) age of onset of learning, b) frequency of communication in the language, and scores on perception and self-reported frequency of use of swearwords in L2, L3, L4 and L5. Quantitative results are supplemented by qualitative data provided by the same participants and elicited by means of open-ended questions.

## 7. Results

This section is divided in three parts. The first part focuses on the effect of type of contact or acquisition context on the two dependent variables using multivariate analyses. The second part presents the results of correlation analyses and multiple linear regression analyses. The third and final part considers the testimonies provided by participants on their experiences with swearing in multiple languages and the effect of type of contact or acquisition context.

### 7.1. The effect of type of contact or acquisition context

Mean scores for language choice for swearing and perceptions of emotional force of swear and taboo words are presented in figure 1. It shows a gradual decrease between L1 and L3 before levelling off towards the L5. Scores for both variables cannot be compared as they are situated on different 5-point scales.



*Figure 1.* Mean scores for frequency of language choice for swearing (FOS) and perceptions of emotional force (PLF) of swear and taboo words in L1, L2, L3, L4 and L5 in the complete database

It was predicted that mixed and naturalistic learning would result in higher levels of sociopragmatic competence compared to instructed learning. To confirm this hypothesis I firstly need to demonstrate that the mixed and naturalistic groups use the L2, L3, L4 and L5 significantly more than the instructed group for swearing (assuming that they will only do so if they

consider it appropriate), and secondly, that scores for emotional force of swear words in these languages are superior for the mixed and naturalistic groups compared to the instructed group. Mean scores for the different languages are presented in figure 2.

#### *7.1.1. L2*

The MANOVAs revealed that type of contact or acquisition context had a significant effect overall on the self-reported use and perceived emotional force of swearwords in the second language (Wilks lambda = .90,  $F = 17.1$ ,  $p < .0001$ ,  $\eta^2 = .054$ ). Scheffé post-hoc tests showed that the instructed group ( $n = 342$ ) scored significantly lower ( $p < .0001$ ) than the mixed ( $n = 409$ ) and the naturalistic ( $n = 140$ ) groups on both frequency of use and perceived emotional force of swearwords. Differences between the mixed group and the naturalistic group were non-significant for both variables (see figure 2). Tests of between-subjects effects revealed that the effect of type of contact or acquisition context was slightly stronger for frequency of use of swearwords ( $F(3) = 22.3$ ,  $p < .0001$ ,  $\eta^2 = .069$ ) compared to perceived emotional force ( $F(3) = 20.6$ ,  $p < .0001$ ,  $\eta^2 = .065$ ).

#### *7.1.2. L3*

The analyses revealed that type of contact or acquisition context had a significant effect overall on the self-reported use and perceived emotional force of swearwords in the third language (Wilks lambda = .89,  $F = 12.7$ ,  $p < .0001$ ,  $\eta^2 = .056$ ). Scheffé post-hoc tests showed that the instructed group ( $n = 410$ ) scored significantly lower than the mixed ( $n = 177$ ) and the naturalistic ( $n = 43$ ) groups on both frequency of use and perceived emotional force of swearwords (with  $p < .0001$  and  $p < .002$  respectively). Differences between the mixed group and the naturalistic group were non-significant for both variables (see figure 2). Tests of between-subjects effects revealed that the effect of type of contact or acquisition context was stronger for frequency of use of swearwords ( $F(3) = 22.1$ ,  $p < .0001$ ,  $\eta^2 = .093$ ) than perceived emotional force ( $F(3) = 20.6$ ,  $p < .0001$ ,  $\eta^2 = .056$ ).

#### *7.1.3. L4*

The analyses revealed that type of contact or acquisition context had a significant effect overall on the self-reported use and perceived emotional force of swearwords in the fourth language (Wilks lambda = .86,  $F = 10.7$ ,

$p < .0001$ ,  $\eta^2 = .073$ ). Scheffé post-hoc tests showed that the instructed group ( $n = 268$ ) scored significantly lower ( $p < .0001$ ) than the mixed group ( $n = 99$ ) but only marginally lower ( $p < .051$ ) than the naturalistic group ( $n = 40$ ) on frequency of use of swearwords. The instructed group did give significantly lower scores for emotional force of swearwords than the mixed group and the naturalistic group ( $p < .0001$ ). Differences between the mixed group and the naturalistic group were non-significant for both variables (see figure 2). Tests of between-subjects effects revealed that the effect of type of contact or acquisition context was equally strong for perceived emotional force ( $F(3) = 21.2$ ,  $p < .0001$ ,  $\eta^2 = .096$ ) and frequency of use of swearwords ( $F(3) = 14.7$ ,  $p < .0001$ ,  $\eta^2 = .095$ ).

#### 7.1.4. L5

The analyses revealed type of contact or acquisition context had a significant effect overall on the self-reported use and perceived emotional force of swearwords in the fifth language (Wilks lambda = .82,  $F = 7.2$ ,  $p < .0001$ ,  $\eta^2 = .092$ ). Scheffé post-hoc tests showed that the instructed group ( $n = 127$ ) scored significantly lower ( $p < .0001$ ) than the mixed group ( $n = 53$ ) but not significantly lower than the naturalistic group ( $n = 33$ ) on frequency of use of swearwords. The instructed group did give significantly lower scores for emotional force of swearwords than the mixed group ( $p < .005$ ) and the naturalistic group ( $p < .024$ ). Differences between the mixed group and the naturalistic group were non-significant for emotional force of

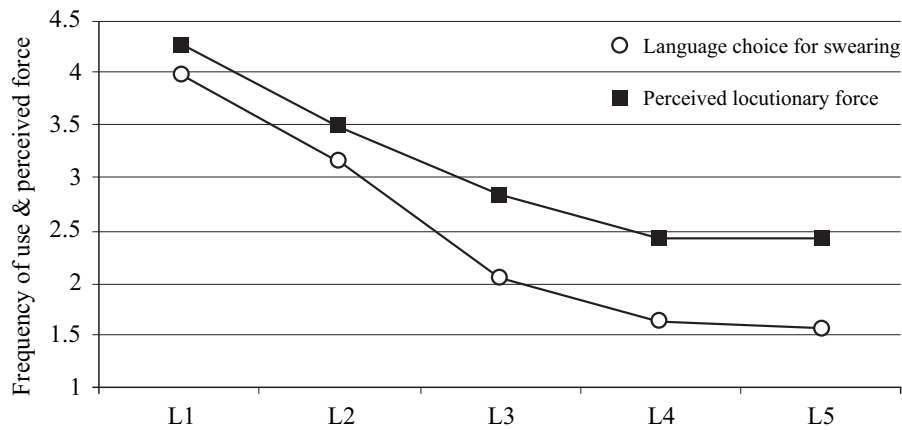


Figure 2. Means for frequency of swearing (FOS) and perceived emotional force (PLF) of swearwords in the L2, L3, L4 and L5

swearwords, but they are significant for frequency of use of swearwords ( $p < .043$ ) (see figure 2). Tests of between-subjects effects revealed that the effect of type of contact or acquisition context was strongest for frequency of use of swearwords ( $F(3) = 12.0$ ,  $p < .0001$ ,  $\eta^2 = .144$ ) and weaker for perceived emotional force ( $F(3) = 6.3$ ,  $p < .0001$ ,  $\eta^2 = .082$ ).

## 7.2. The effects of age of onset and frequency of use of the languages

The statistical analyses showed consistently significant effects for type of contact or acquisition context on perception and self-reported use of swearwords. Nevertheless, the values of  $\eta^2$ , a measure of effect strength, remained relatively modest, varying typically between 5% and 10%. Other independent variables may be linked to the two dependent variables. The participants supplied information concerning age of onset of learning the different languages as well as frequency of use of these languages in the first part of the questionnaire. Research into the effect of age of onset on ultimate attainment has shown that the younger one is exposed to a language the higher the probability of reaching high levels of proficiency (e.g. DeKeyser 2000). Learners who started the learning process under age 12 seem to have an advantage, although this does not mean that some older learners could not reach high levels of proficiency. Singleton (2001) however points out that it is always difficult to know whether the cause of the observed difference is age of onset of learning or the much longer duration of exposure. Recent studies on the effect of age of onset do suggest that learners who started using implicit learning mechanisms at a younger age generally outperform those who embarked on the language learning process at a later age and use alternative mechanisms linked to declarative memory systems (DeKeyser 2000; Romero Trillo 2002; Ullman 2001). Perception of emotional force probably relies on implicit knowledge rather than explicit knowledge. A negative correlation could therefore be expected between age of onset of learning and perceived emotional force of swearwords.

Frequency of speaking the language has also been linked to higher levels of sociopragmatic competence. Speakers who came back from periods of study abroad used higher proportions of informal syntactical, lexical, morphological and phonological variants (Dewaele and Regan 2001, 2002; Mougeon et al. 2002; Rehner et al. 2003; Sax 2003). A positive correlation can therefore be expected between general frequency of use of the language and the dependent variables. The results of the correlation analysis are presented in table 1.

*Table 1.* Pearson correlations between age of onset of learning, frequency of use of the TL and measures of perception of emotional force and self-reported language choice for swearing

	L2		L3		L4		L5	
	Age of onset	Frequency of use	Age of onset	Frequency of use	Age of onset	Frequency of use	Age of onset	Frequency of use
Language choice for swearing	-0.10***	0.46***	-0.06	0.45***	-0.05	0.49***	-0.06	0.53***
Emotional force	-0.13***	0.26***	-0.15***	0.31***	0.06	0.36***	0.00	0.42***

p < .05 \*\* p < .001 \*\*\* p < .0001

The analyses show that age of onset is significantly negatively correlated with the dependent variables for the L2, but the relation weakens in the L3 and disappears in the L4 and L5. In other words, the lower the age of onset of learning the L2 and L3, the higher the scores on self-reported use of the TL for swearing and perception of emotional force (but not so for the L4 and L5). It thus shows that the age effect is most significant for the L2 and L3. This could be related to the fact that the average age at which the learning of the L2 was undertaken was significantly lower ( $t(892) = -23.7, p < .0001$ ) than the L3 which in turn was significantly lower ( $t(630) = -17.9, p < .0001$ ) than the L4 and again significantly lower ( $t(341) = -13.5, p < .0001$ ) than the L5 (see table 2).

*Table 2.* Age of onset of language learning

Language	Average age	Standard deviation
L2	8.5	6.3
L3	13.6	6.6
L4	17.7	6.8
L5	21.9	8.1

Correlation values between frequency of use of the language and the dependent variables are even more significant (with  $p < .0001$ ). This seems logical as authentic usage of the language (in combination with instruction) is generally regarded as the best way to develop one's grammatical and

sociopragmatic skills (Byram 1997). Ellis (2002) argued that learners' inter-languages are derived from input frequencies. As the probability of hearing swearwords in the TL outside the language class is much greater than within the walls of the classroom, one can expect these to be (un)consciously noticed and linked to a specific social context. This noticing results in the "selective internalization of language input in interaction with various L2 speakers" (Tarone 2002: 287).

Linear multiple regression analyses were performed in order to identify the best predictors for language choice in swearing and perceived emotional force of swearwords in speaking L2, L3, L4 and L5. The multiple regression analysis revealed a significant effect of age of onset and frequency of use of the language on frequency of choice of a particular language for swearing. This is the case for the L2 ( $R^2 = .22$ , frequency of use:  $\beta = .46$ ,  $t = 13.3$ ,  $p < .0001$ ; age of onset:  $\beta = -.09$ ,  $t = -2.6$ ,  $p < .009$ ); for the L3 ( $R^2 = .21$ , frequency of use:  $\beta = .46$ ,  $t = 13.4$ ,  $p < .0001$ ; age of onset:  $\beta = -.02$ ,  $t = -.43$ ,  $p = ns$ ); for the L4 ( $R^2 = .25$ , frequency of use:  $\beta = .51$ ,  $t = 10.7$ ,  $p < .0001$ ; age of onset:  $\beta = -.09$ ,  $t = -1.9$ ,  $p < .06$ ); for the L5 ( $R^2 = .32$ , frequency of use:  $\beta = .56$ ,  $t = 8.9$ ,  $p < .0001$ ; age of onset:  $\beta = -.16$ ,  $t = -2.5$ ,  $p < .015$ ).

A second series of multiple regression analyses was performed in order to determine the effect of frequency of use and age of onset on the perceived emotional force of swearwords in the different languages. As before, all the analyses yielded significant results, with frequency of use appearing to be the strongest predictor and age of onset gradually weakening in languages learnt later in life. These are the results: for the L2: ( $R^2 = .09$ , frequency of use:  $\beta = .26$ ,  $t = 6.9$ ,  $p < .0001$ ; age of onset:  $\beta = -.15$ ,  $t = -4.0$ ,  $p < .0001$ ); for the L3: ( $R^2 = .11$ , frequency of use:  $\beta = .30$ ,  $t = 7.0$ ,  $p < .0001$ ; age of onset:  $\beta = -.11$ ,  $t = -2.4$ ,  $p < .015$ ); for the L4: ( $R^2 = .13$ , frequency of use:  $\beta = .36$ ,  $t = 6.8$ ,  $p < .0001$ ; age of onset:  $\beta = .01$ ,  $t = -.14$ ,  $p = ns$ ); for the L5: ( $R^2 = .17$ , frequency of use:  $\beta = .41$ ,  $t = 5.6$ ,  $p < .0001$ ; age of onset:  $\beta = .02$ ,  $t = -.26$ ,  $p = ns$ ).

It thus seems that age of onset and frequency of use of a language have the strongest effect on language choice for swearwords, and have a weaker, but still significant, effect on perception of emotional force of these swearwords.

### 7.3. Testimonies by participants

The results of the quantitative analysis are backed up by the information provided by the participants. Some multilinguals are acutely aware of the

incompleteness of their sociopragmatic competence in the TL learnt at school, and the resulting difficulties in authentic communication in the TL:

Bart, a 24 year-old-male (Dutch L1, French L2, English L3) and instructed user of French writes: in school we learn how to use French in a polite and friendly way but when I am calling to the Customer Service of a French company to complain about something and want to sound a bit more severe irritated angry ... then it is difficult to find that severe irritated angry tone because you are concentrating on French grammar and vocabulary... I wouldn't have to do that in Dutch.

Some multilinguals who teach their native language in a foreign country noticed that their L1 becomes more artificial in the classroom. They somehow feel that the use of their usual vernacular style including colloquial vocabulary would be inappropriate. This "teacher talk" seems to appear automatically reports Johanna, a 27-year-old female (English L1, French L2, Italian L3, Spanish L4): As an English teacher I also use my mother tongue on a daily basis in a completely non-emotional and artificial context. Somehow I don't really count my class English as English though because it's not a normal use of language: I have to carefully monitor my vocabulary level, adopt a neutral non-American accent that I would feel ridiculous using with native speakers (of any nationality!) and break into Italian for grammatical explanations.

One participant, Henry (22-year-old male, English L1, French L2, Italian L3), currently studying French, reported that after having spent his year abroad in France, he felt bored in the French language class: My French is rather good now. I can speak as they do in France. But I'm not allowed to speak that way here during classes. We're all supposed to use a formal speech style and I find that so artificial.

The following examples show that context of acquisition and type of contact with the TL is but one factor in the complex interplay of variables that determine perception and use of swearwords. Dewaele (2004b) pointed out that depending on the intended illocutionary effects and the identity of the interlocutor, speakers can either choose to swear in a language in which they can express themselves forcefully, or choose a language in which they perceive the swearwords to be weaker.

Jessamy, a 34 year-old female (English L1, Spanish L2 – naturalistic) writes: We speak in English. We argue in English though I will scream out Spanish swear words to vent my frustration. This of course doesn't help matters as my partner cannot speak Spanish.

A 53 year-old female (Cantonese L1, English L2 – instructed, French L3, Putonghua L4, Japanese L5) from Hong Kong who preferred to remain



anonymous, reports on the social stigma attached to swearing in her mother tongue, hence her preference for English: I find it more difficult to swear in Cantonese than in English. Swearing in Cantonese is a big taboo for people of my educational level however swearing in English doesn't sound vulgar to me though it takes off the emotional intensity.

## **8. Discussion and conclusion**

This study has shown that type of contact or acquisition context has a significant effect on aspects of sociopragmatic competence in consecutive non-native languages, but that other independent variables, such as age of onset of learning a language and frequency of use of the language, have an even stronger effect. The findings of the study fully support the first hypothesis, viz. that mixed and naturalistic language users score higher on self-reported use and perceived emotional force of swearwords than instructed language users. The findings also fully support the second hypothesis, that age of onset and frequency of use strongly predict scores on the self-reported use and perceived emotional force of swearwords.

The re-examination of the effects of instructed versus mixed contexts confirm the superiority of mixed learning versus instructed learning (cf. Housen 2002). The use of MANOVAs and multiple linear regression analyses allowed me to measure the exact weight of the effect of instruction and type of contact (as well as age of onset of learning and frequency of use) on every language acquired by 1039 multilinguals after a period that ranged from zero to almost 70 years. The patterns that emerged were remarkably consistent across languages (L2-L5) and across individuals. They were strengthened by individual testimonies of participants relating their learning experiences.

It is not surprising that if one's contact with a TL has been limited to the classroom the resulting communicative competence will be more limited compared to those users who have experienced and used the TL in wider variety of situations. And yet, although the effect of the learning context and type of contact was found to be consistently significant, it is not the strongest predictor for the dependent variables. The indices of effect size showed relatively modest values for the different languages. Subsequent correlation analyses and multiple linear regressions revealed that age of onset of learning, and, to a stronger extent, frequency of use of the language accounted for a larger proportion of the variance in the data. The dependent variables were also found to correlate with each other. This suggests that

the use of swearwords is linked to the perception of strength one has of these words. It is impossible, however, to know whether perception precedes use, or whether use results in native-like perception.

The findings have three important pedagogical implications. Firstly, instructed learning should ideally rely on a rich source of diverse types of written and visual authentic material (Ellis 1999; Weyers 1999) allowing learners to familiarise themselves with sociolinguistic, sociopragmatic and sociocultural aspects of the TL. Secondly, formal instruction should be complemented by a period in the TL community in order to stimulate the process of “temporary re-socialisation into a foreign culture and its practices and beliefs” (cf. Alred and Byram 2002: 340) hence my strong support for study abroad programs. Thirdly, whatever type of language instruction/contact one opts for, the results suggest that the younger one starts it, the better the end result will be.

### **Acknowledgements**

I would like to thank Aneta Pavlenko, the volume editors and the anonymous reviewer for their excellent comments on a previous version of this paper.

### **References**

- Alred, Geoff and Mike Byram  
 2002      Becoming an intercultural mediator: A longitudinal study of residence abroad. *Journal of Multilingual and Multicultural Development* 23 (5): 339–352.
- Adamson, Hugh D.  
 1988      *Variation theory and second language acquisition*. Washington: Georgetown University Press.
- Adamson, Hugh D. and Vera Regan  
 1991      The acquisition of community speech norms by Asian immigrants learning English as a second language. *Studies in Second Language Acquisition* 13: 1–22.
- Austin, John Langshaw  
 1962      *How to do things with words*. Oxford: Oxford University Press.

- Bayley, Robert  
1991 Variation theory and second language learning: Linguistic and social constraints in interlanguage tense marking. Ph.D. diss., Stanford University.
- Beebe, Lesley  
1988 Five sociolinguistic approaches to Second Language Acquisition. In *Issues in Second Language Acquisition: Multiple Perspectives*, L. Beebe (ed.), 43–75. Cambridge, Mass.: Newbury House.
- Blondeau, Hélène, Naomi Nagy, Gillian Sankoff and Pierrette Thibault  
2002 La couleur locale du français L2 des anglo-montréalais. In *Appropriation de la variation en français langue étrangère. Special issue of AILE 17*, Jean-Marc Dewaele and Raymond Mougeon (eds.), 73–100.
- Byram, Mike  
1997 Teaching and Assessing Intercultural Communicative competence. Clevedon: Multilingual Matters.
- Chavez, Monika  
1998 Learner's perspective on authenticity. *International Review of Applied Linguistics* 36 (4): 277–306.
- Coleman, James A.  
1996 *Studying languages: A survey of British and European students, the proficiency, background, attitudes and motivations of students of foreign languages in the United Kingdom and Europe*. London: Centre for Information on Language Teaching and Research.
- Cook, Vivian (ed.)  
2002 *Portraits of the L2 User*. Clevedon: Multilingual Matters.
- Critchley, Sonia  
1994 Discourse variety in contemporary French language: A pedagogical approach. In *Discourse variety in contemporary French: Descriptive and pedagogical approaches*, James A. Coleman and Robert Crawshaw (eds.), 203–236. London: CILT.
- Cuq, Jean-Pierre  
1994 Cadre théorique et modalités d'insertion de la variété linguistique dans l'enseignement du français langue étrangère et seconde. In *Discourse variety in contemporary French: Descriptive and pedagogical approaches*, James A. Coleman and Robert Crawshaw (eds.), 19–34. London: CILT.
- Davies, Alan  
2003 *The Speaker: Myth and Reality*. Native Clevedon: Multilingual Matters.
- DeKeyser, Robert  
2000 The robustness of Critical Period effects in Second Language Acquisition. *Studies in Second Language Acquisition* 22(4): 499–533.

## Dewaele, Jean-Marc

- 1999 Word order variation in interrogative structures of native and non-native French. *ITL Review of Applied Linguistics* 123/124: 161–180.
- 2004a The emotional force of swearwords and taboo words in the speech of multilinguals. In *Language and Emotions: A Crosslinguistic Perspective. Special issue of Journal of Multilingual and Multicultural Development* 25 (2–3), Aneta Pavlenko and Jean-Marc Dewaele (eds.), 204–219.
- 2004b Blistering barnacles! What language do multilinguals swear in?! In *Bilingualism and Emotions. Special issue of Estudios de Sociolingüística* 5 (1), Aneta Pavlenko and Jean-Marc Dewaele (eds.) 83–105.

## Dewaele, Jean-Marc and Aneta Pavlenko

- 2001 *Web questionnaire Bilingualism and Emotions*. University of London.
- 2002 Emotion vocabulary in interlanguage. *Language Learning* 52 (2): 265–324.

## Dewaele, Jean-Marc and Vera Regan

- 2001 The use of colloquial words in advanced French interlanguage. In *EUROSLA Yearbook 1*, Susan Foster-Cohen and Anna Nizegorodcew (eds.), 51–68. Amsterdam/Philadelphia: Benjamins.
- 2002 Maîtriser la norme sociolinguistique en interlangue française: le cas de l'omission variable de 'ne'. *Journal of French Language Studies* 12: 123–148.

## Dewaele, Jean-Marc and Nathalie Wourm

- 2002 L'acquisition de la compétence sociopragmatique en langue étrangère. *Revue Française de Linguistique Appliquée* 7 (2): 129–143.

## Dörnyei, Zoltan

- 2003 *Questionnaires in second language research: Construction, administration, and processing*. Mahwah, NJ: Lawrence Erlbaum.

## Drescher, Martine

- 2000 Eh tabarnouche! c'était bon. Pour une approche communicative des jurons en français québécois. *Cahiers de praxématique* 34: 133–160.

## Dufour, Robert and Judith Kroll

- 1995 Matching words to concepts in two languages: A test of the concept mediation model of bilingual representation. *Memory and Cognition* 23: 166–180.

## Ellis, Rod

- 1989 Are classroom and naturalistic acquisition the same? A study of the classroom acquisition of German word order rules. *Studies in Second Language Acquisition* 11: 305–328.
- 1990 *Instructed Second Language Acquisition: Learning in the Classroom*. Oxford: Blackwell.
- 1992 Learning to communicate in the classroom: A study of two language learners' requests. *Studies in Second Language Acquisition* 14: 1–23.

- 1999 Retrospect and Prospect. In *Learning A Second Language Through Interaction*, R. Ellis (ed.), 233–257. Amsterdam/Philadelphia: Benjamins.
- 2002 Frequency effects in language processing. A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition* 24: 223–236.
- Evans, Colin
- 1988 *Language People. The experience of teaching and learning Modern Languages in British Universities*. Milton Keynes/Philadelphia: Open University Press.
- Freed, Barbara F. (ed.)
- 1995 *Second Language Acquisition in a Study Abroad Context*. Amsterdam/Philadelphia: Benjamins.
- Giles, Howard and Nicolas Coupland
- 1991 *Language contexts and consequences*. Milton Keynes: Open University Press.
- Gonzalez-Reigosa, Fernando
- 1976 The anxiety arousing effect of taboo words in bilinguals. In *Cross-cultural anxiety*, Charles D. Spielberger and Rogelio Diaz-Guerrero (eds.), 89–105. Washington, D.C.: Hemisphere.
- Graham, C. Ray, Arien W. Hamblin and Stanley Feldstein
- 2001 Recognition of emotion in English voices by speakers of Japanese, Spanish and English. *International Review of Applied Linguistics* 39: 19–37.
- Harley, Birgit, Paul Allen, Jim Cummins and Merrill Swain (eds.)
- 1990 *The development of second language proficiency*. Cambridge: Cambridge University Press.
- Harris, Catherine L., Ayse Ayçiçeği and Jean B. Gleason
- 2003 Taboo words and reprimands elicit greater autonomic reactivity in a first than in a second language. *Applied Psycholinguistics* 24: 561–579.
- Hoffman Hicks, Sheila
- 2000 The longitudinal development of French foreign language pragmatic competence: Evidence from study abroad participants. PhD diss., Indiana University.
- House, Juliane
- 1996 Developing pragmatic fluency in English as a Foreign Language: Routines and metapragmatic awareness. *Studies in Second Language Acquisition* 18 (2): 225–252.
- Housen, Alex
- 2002 Paramètres liés à l'environnement dans l'apprentissage guidé d'une langue seconde. In *L'acquisition en classe de langue. Special issue of AILE 16*, Daniel Coste (ed.) 23–46.

- Hymes, Dell  
 1979 Reprint. On communicative competence. In *The communicative approach to language teaching*, Christopher J. Brumfit and Keith Johnson (eds.), 5–26. Oxford: Oxford University Press, 1970
- Kasper, Gabriele  
 2001 Four perspectives on L2 pragmatic development. *Applied Linguistics* 22 (4): 502–530.
- Kasper, Gabriele and Kenneth R. Rose  
 1999 Pragmatics and second language acquisition. *Annual Review of Applied Linguistics* 19: 81–104.  
 2001 Pragmatics in Language Teaching. In *Pragmatics in Language Teaching*, Kenneth R. Rose and Gabriele Kasper (eds.), 1–9. Cambridge: Cambridge University Press.  
 2002 Pragmatic development in a second language. *Language Learning* 52 supplement 1, 1–352.
- Kecskes, Istvan  
 1998 The state of L1 knowledge in foreign language learners. *Word* 49: 321–341.
- Kecskes, Istvan and Tunde Papp  
 2000 *Foreign language and mother tongue*. Hillsdale, NJ: Lawrence Erlbaum.
- Kroll, Judith, Erica Michael, Natasha Tokowicz and Robert Dufour  
 2002 The development of lexical fluency in a second language. *Second Language Research* 18: 141–175.
- Lapkin, Sharon, Doug Hart and Merrill Swain  
 1995 A Canadian interprovincial exchange: evaluating the linguistic impact of a three-month stay in Quebec. In *Second Language Acquisition in a Study Abroad Context*, Barbara F. Freed (ed.), 67–94. Philadelphia: Benjamins.
- Leech, Geoffrey  
 1983 *Principles of Pragmatics*. London: Longman.
- Lyster, Roy  
 1994 The effect of functional-analytic teaching on aspects of French immersion students' sociolinguistic competence. *Applied Linguistics* 15: 263–287.  
 1996 Question forms, conditionals, and second-person pronouns used by adolescent native speakers across two levels of formality in written and spoken French. *Modern Language Journal* 80: 165–180.  
 forthc. Classroom SLA and pedagogical implications for French as a second/foreign language. *Journal of French Language Studies* 14.
- Lyster, Roy and Jacques Rebuffot  
 2002 Acquisition des pronoms d'allocution en classe de français immersif. In *Appropriation de la variation en français langue étrangère. Special issue of AILE 17* Jean-Marc Dewaele and Raymond Mougeon (eds.), 51–72.

- Mougeon, Raymond, Terry Nadasdi and Katherine Rehner  
2002 État de la recherche sur l'appropriation de la variation par les apprenants avancés du FL2 ou FLE. In *Appropriation de la variation en français langue étrangère. Special issue of AILE 17*, Jean-Marc Dewaele and Raymond Mougeon (eds.), 7–50.
- Myers Scotton, Carol and Janice Bernstein  
1988 Natural conversations as a model for textbook dialogue. *Applied Linguistics* 9 (4): 372–384.
- Offord, Malcolm  
1994 Teaching varieties and register. In *Discourse variety in contemporary French: Descriptive and pedagogical approaches*, James A. Coleman and Robert Crawshaw (eds.), 185–202. London: CILT
- Pavlenko, Aneta  
1999 New approaches to concepts in bilingual memory. *Bilingualism: Language and Cognition* 2: 209–230.  
2002 Poststructuralist approaches to the study of social factors in L2. In *Portraits of the L2 User*, Vivian Cook (ed.), 277–302. Clevedon: Multilingual Matters.
- Pica, Teresa  
1983 Adult acquisition of English as a second language under different conditions of exposure. *Language Learning* 33: 465–97.  
1995 The acquisition of sociolinguistic native speech norms: Effects of a year abroad on L2 learners of French. In *Second Language Acquisition in a Study Abroad Context*, Barbara Freed (ed.), 245–267. Amsterdam/Philadelphia: Benjamins.  
1996 Variation in French interlanguage: A longitudinal study of sociolinguistic competence. In *Language Acquisition and Linguistic Variation*, Robert Bayley and Dennis Preston (eds.), Second 177–201. Amsterdam/Philadelphia: Benjamins.  
2002 Le contexte d'acquisition: la variation du groupe et de l'individu. In *Appropriation de la variation en français langue étrangère. Special issue of AILE 17*, Jean-Marc Dewaele and Raymond Mougeon (eds.), 123–141.
- Regan, Vera  
2004 From speech community back to classroom: What variation analysis can tell us about the role of context in the acquisition of French as a foreign language. In *Focus on French as a Foreign Language: Multidisciplinary Approaches*, Jean-Marc Dewaele (ed.), 191–209. Clevedon: Multilingual Matters.
- Rehner, Katherine and Raymond Mougeon  
1999 Variation in the spoken French of immersion students: To ne or not to ne, that is the sociolinguistic question. *Canadian Modern Language Review* 56: 124–154.

- Rehner, Katherine, Raymond Mougeon and Terry Nadasdi  
 2003 The learning of sociolinguistic variation by advanced FSL learners: The case of *nous* versus *on* in immersion French. *Studies in Second Language Acquisition* 25 (1): 127–157.
- Rintell, Ellen  
 1984 But how did you feel about that? The learner's perception of emotion in speech. *Applied Linguistics* 5: 255–264.
- Romero Trillo, Jesus  
 2002 The pragmatic fossilization of discourse markers in non-native speakers of English. *Journal of Pragmatics* 34: 769–784.
- Sankoff, Gillian, Pierrette Thibault, Naomi Nagy, H el ene Blondeau, Marie Odile Fonollosa and Lucille Gagnon  
 1997 Variation in the use of discourse markers in a language contact situation. *Language Variation and Change* 9 (2): 191–218.
- Sax, Kelly  
 2003 Acquisition of Stylistic Variation in American Learners of French. Ph.D. diss., Indiana University.
- Sercu, Lies  
 2000 *Acquiring Intercultural Communicative Competence from Textbooks*. Leuven: Leuven University Press.
- Singleton, David  
 2001 Age and Second Language Acquisition. *Annual Review of Applied Linguistics* 21: 77–89.
- Tarone, Elaine  
 1988 *Variation in interlanguage*. London: Edward Arnold.  
 1990 On variation in interlanguage: A response to Gregg. *Applied Linguistics* 11: 392–400.  
 1998 A Sociolinguistic Perspective on an SLA Theory of Mind. *Studia Anglica Posnaniensa* 33: 431–444.  
 2002 Frequency effects, noticing and creativity. Factors in a variationist interlanguage framework. *Studies in Second Language Acquisition* 24: 287–296.
- Tarone, Elaine and Merrill Swain  
 1995 A sociolinguistic perspective on second language use in immersion classrooms. *Modern Language Journal* 79: 166–178.
- Thomas, Alain  
 2002 La variation phon etique en fran ais langue seconde au niveau universitaire avanc e. In *Appropriation de la variation en fran ais langue  trang ere*. Special issue of *AILE* 17, Jean-Marc Dewaele et Raymond Mougeon (eds.), 101–121.
- Toya, Mitsuyo and Mary Kodis  
 1996 But I don't want to be rude: On learning how to express anger in the L2. *JALT Journal* 18 (2): 279–295.



- Towell, Richard and Jean-Marc Dewaele  
2004 The role of psycholinguistic factors in the development of fluency amongst advanced learners of French. In *Focus on French as a Foreign Language: Multidisciplinary Approaches*, Jean-Marc Dewaele (ed.), 210–239. Clevedon: Multilingual Matters.
- Towell, Richard, Hawkins, Roger and Nives Bazergui  
1996 The development of fluency in advanced learners of French. *Applied Linguistics* 17 (1): 84–119.
- Ullman, Michael T.  
2001 The neural basis of lexicon and grammar in first and second language: the declarative/procedural model. *Bilingualism: Language and Cognition* 4 (1): 105–122.
- Valdman, Albert  
2003 The acquisition of sociostylistic and sociopragmatic variation by instructed second language learners: The elaboration of pedagogical norms. In *The Sociolinguistics of Foreign Language Classrooms: Contributions of the Native, the Near-Native, and the Non-Native Speaker*, Carl Blyth (ed.), 57–78. Boston: Heinle Thomson.
- Weyers, Joseph R.  
1999 The effect of authentic video on communicative competence. *The Modern Language Journal* 83: 339–349.
- Young, Richard  
1991 *Variation in Interlanguage Morphology*. New York: Peter Lang.  
1999 Sociolinguistic approaches to SLA. *Annual Review of Applied Linguistics* 19: 105–32.



## Index

- Accuracy, 13, 19, 21, 41, 54, 58–70, 83, 87, 89, 99–123, 132, 183, 200, 215, 216, 220, 249, 252–254, 259–262, 273, 286, 294, 333, 335, 338, 339, 348, 374, 397, 412, 413, 424, 426, 428, 437, 439, 452, 461, 498–502, 512, 516–520, 535
- acquisition (learning) context, 21, 531, 533, 540–547, 551
- advanced learner, 19, 21, 34, 52, 159, 383–385, 393–399, 495, 499, 500, 503, 506, 519, 524, 536, 538
- adverb, 17, 132, 201, 292–301, 365, 384, 389, 464
- age, 4, 9, 22, 120, 121, 347, 539–543, 547–551
- Analytic ability, 13, 99–110, 118, 120
- applied linguistics, 271–274, 540
- Aptitude, 9, 101–105, 119–121, 287, 451, 485
- Arab(ic), 15, 123, 167, 169, 175–177, 181, 189–194
- Attention, 6, 31, 44, 103, 123, 158, 191, 200, 238, 243, 261, 288–284, 313, 321, 323, 333, 337, 355, 383, 408, 411, 424, 426, 428, 437, 440–442, 477
- Audiolingual(ism), 17, 102, 273, 284, 285
- Automatisation (automaticity/automatic), 3, 6, 8, 13, 31, 40–47, 103, 106, 120, 138, 139, 283, 300, 414, 441
- Awareness, 6–9, 19, 123, 133, 199, 217, 218, 229, 261, 274, 297, 314, 341, 355, 356, 359, 420, 423, 534, 538
- Categorization, 64, 65, 70, 74, 84, 534
- chunking, 103, 442, 443, 446, 451
- clausal planning, 13, 31, 39, 41, 47
- Content and Language Integrated Learning (CLIL), 20, 433–452
- Cognitive style, 9
- collaborative dialogue / discourse, 19, 353, 356, 372, 395, 397, 483
- communicative language teaching / method / approach, 100, 102, 121, 200, 225–228, 272–301, 285, 338, 397, 400, 407, 434, 438, 439, 503, 534
- competence, 5, 12, 236, 239, 445–447, 538–540
- communicative, 20, 458, 460, 531–535, 540, 551
- discourse, 21, 457–459, 458, 480, 482, 484
- grammatical, 21, 297, 485, 502, 540
- lexical, 500
- linguistic, 40, 44, 46, 458, 533
- sociocultural, 533, 538–540
- sociolinguistic, 104, 500, 501, 535–538
- sociopragmatic, 22, 538, 544, 547, 550, 551
- strategic, 503
- Competition Model, 13, 64
- “complexity (structural; rule)”, 11, 16, 32, 53, 104, 106, 117, 119, 131, 136–139, 142, 152, 155, 226, 235–264, 358, 374, 464, 465, 470, 472–479, 498, 499, 535
- connectionist, 54, 76–80, 85, 87, 92, 93, 134, 440, 441
- Consciousness, 6, 23, 132
- consciousness raising, 3

- content-based instruction / learning /  
 approach / syllabus, 132, 200, 202,  
 339, 433–439, 442
- Context analysis, 461, 464, 465, 481
- contrastive analysis, 273, 275, 458
- Controlled practice, 10, 14, 15, 106, 246
- corrective feedback, 18, 53, 79, 82, 200,  
 201, 209, 211, 281–288, 298, 300,  
 333–349, 459, 463, 477
- Conversational analysis, 461
- Cue-category-mapping, 64, 65, 73, 87
- decision making task, 386, 387
- default hypothesis of vocabulary acqui-  
 sition, 17, 311–314, 318–320, 322,  
 324
- Definite article, 13, 53, 56, 58, 72, 76, 85,  
 112, 203, 204
- dictogloss task, 18, 19, 354, 358, 359,  
 362, 375, 398
- discourse acquisition, 458–460, 483, 484
- discourse analysis, 338, 461, 463, 465,  
 474, 475, 481
- Dual mechanism/system model, 131–  
 135, 155, 156, 411
- Emotional force, 531, 532, 539, 543–554
- End-state of acquisition/learning, 12, 19
- English, 12–20, 32, 35, 39, 58, 99–123,  
 132, 136, 167–195, 199–229, 288–  
 293, 325, 336, 338, 347, 353,  
 354, 359, 383–388, 408–429, 438,  
 442, 445, 452, 509, 535–538, 550,  
 551
- Error correction, 9, 18, 109, 336, 338
- ESL, 13, 14, 99, 100, 106–112, 118, 121,  
 124, 132, 133, 169–172, 199, 201,  
 208–217, 224, 228, 229, 280,  
 284–287
- explicit  
 learning/acquisition, 7, 33, 411, 445  
 instruction/teaching, 16, 20, 131–137,  
 192, 201, 202, 205, 208, 209, 225,  
 226, 228, 235–238, 243, 248–250,  
 254, 256, 260–263, 282, 298, 411,  
 534
- feedback  
 corrective, 18, 53, 79, 82, 199, 201,  
 209, 281, 282, 285, 288, 298,  
 300, 333–340, 346, 347, 459,  
 463, 477  
 negative, 18, 158, 333–338, 342, 344,  
 415, 419
- fluency, 12–14, 19, 31–47, 99–108, 113,  
 116–124, 283, 412, 413, 439, 441,  
 499–501, 526, 535
- Focus on Form (FoF), 3, 17–19, 124,  
 132–134, 156, 271, 272, 277–279,  
 281–287, 295, 311, 321–324, 334,  
 335, 344, 354, 355, 394, 409
- Focus on Forms (FoFs), 15, 17, 18, 132,  
 271, 277, 278, 286, 287, 295, 311,  
 321–324, 334, 339, 344, 348, 354
- Foreigner talk, 459
- Form-meaning mapping, 15, 16, 199, 200,  
 441
- Form-function mapping, 9, 239, 242, 423
- French, 12, 16–21, 31–45, 54, 55, 71, 76,  
 104–109, 133, 159, 175, 176, 189–  
 192, 199–229, 235, 239–260, 282,  
 290–293, 338, 408–428, 438, 457–  
 464, 480, 495–523, 532, 536–538,  
 541, 550
- Frequency (of use/effects), 22, 140, 141,  
 159, 353, 514, 531, 533, 535, 539,  
 541, 543, 545–551
- input, 63, 70, 87, 131, 155
- token, 13, 69, 70, 158
- type, 13, 14, 69, 140–142, 147, 148,  
 152–155, 158, 159
- Gender, 21, 51–93, 112, 113, 121, 203,  
 204, 219, 228
- German, 13, 18, 52–91, 136, 333–339,  
 388, 438, 499

- Grammar  
 instruction/teaching, 17, 32, 40, 202, 235, 237, 261, 271–284, 322, 344  
 rules, 14, 133, 414, 419
- Grammaticality judgement, 8, 16, 133, 169, 171, 177–182, 186, 187, 193, 195, 214, 221, 230, 235, 246–250, 254–256, 293–297, 433, 442–450
- imitation, 442, 445–448
- immersion, 18, 104, 133, 200, 335, 339, 342, 344–348, 434–439, 452, 458, 523, 535, 536, 542
- Imparfait, 241, 479, 491, 505
- Implicational generalization hypothesis (IGH), 15, 167, 169, 175
- implicit learning, 20, 440, 533, 547
- Incidental  
 acquisition/learning, 23, 272, 280, 281, 319, 320, 323, 534  
 instruction/teaching/focus, 312, 322, 324, 325
- information gap task, 386, 387
- information processing model, 354, 355–358, 373
- Input  
 classroom, 13, 14, 459, 536  
 comprehensible, 4, 9, 132, 277, 288, 335, 399, 408, 439, 440, 450  
 enhancement, 3, 9, 87, 132  
 flooding, 3, 9  
 frequency, 63, 70, 87, 131, 155  
 Hypothesis, 9, 408  
 modified, 357–359, 372, 384, 395, 445  
 processing, 4, 354
- Instruction  
 Communication-Focused, 9, 14, 18, 331  
 contrastive, 211, 225  
 Form-Focused, 9, 10, 14–17, 99, 119, 120, 132, 165, 167, 199–228, 225, 227, 228, 254, 277, 279, 301, 320, 321, 397, 534  
 metalinguistic, 35, 102, 226  
 planned lexical, 17, 18, 311, 321–324
- Intake, 7, 92, 132, 133, 335
- Intentional learning, 7, 23, 40, 323, 411, 542
- interaction  
 classroom, 209, 334, 336, 460, 465, 474, 482–484  
 Hypothesis, 228, 277, 279–283, 408  
 modified, 334, 385, 463, 470–474
- Interrogatives, 17, 288, 301
- Japanese, 13, 57, 58, 80, 314, 400, 427, 461, 478, 485, 498, 500, 537
- Knowledge  
 analysed, 7, 8  
 conscious, 1, 7, 40  
 declarative, 3, 8, 44, 106, 299, 388, 414, 421, 441  
 explicit, 7, 8, 40, 41, 44–46, 239, 248, 262, 373, 411, 500, 534, 547  
 implicit, 7, 8, 16, 40, 44–46, 235, 236, 248, 254, 261, 262, 372, 373, 411, 445, 534, 547  
 metalinguistic, 3, 7, 40, 44, 46, 236, 485, 503  
 procedural, 7, 12, 19, 106, 262, 414, 441  
 subconscious, 1, 20
- Markedness, 5, 11, 15, 16, 167–169, 235, 242–244, 249, 254, 260–262
- memory, 31, 40, 101, 105, 110, 133, 440–446  
 associative, 134, 135  
 explicit, 46, 248  
 procedural, 20, 40, 45, 46, 485  
 declarative/semantic, 40, 44, 46, 485, 537, 547  
 working, 43–46, 103, 357–359, 415, 441  
 short term, 41, 441, 443, 446, 451  
 long term, 41, 45, 313, 357–359, 441

- metacognition, 18, 354–356
- metalinguistic  
 activity/task, 13, 19, 99–101, 131, 212, 213, 228, 411–413, 423, 428  
 awareness, 199, 229, 355, 423, 534  
 comments, 341, 416  
 feedback, 18, 341–346  
 rules, 9, 11, 238
- Monitor(ing), 1, 3, 8, 32, 34, 42–44, 415, 550
- Morphological processing, 14, 131, 134–136
- Motivation, 4, 6, 9, 134, 226, 287, 387, 408, 451, 460, 462, 475, 482, 503
- Natural approach, 17
- Negation, 16, 235, 244, 245, 254–256, 261, 536
- Negotiation of meaning/form, 9, 18, 19, 108, 201, 226, 334–336, 340, 344, 347, 348, 375, 383–385, 399, 400, 408, 409, 452, 459, 482, 483
- Network Model, 13, 70, 87
- nonce word, 58–62, 71
- noticing, 7, 9, 18, 19, 102, 103, 120, 132, 242, 261, 311–315, 323, 324, 333, 335, 338, 344, 347, 358–360, 364–375, 441, 442, 448, 549
- Noun, 12, 32–40, 60–91, 113, 117, 167–191, 199–227, 242
- Noun phrase accessibility hypothesis (NPAH), 168, 171, 174, 178, 188, 192
- Order of acquisition/learning/development, 64, 65, 75, 541
- Output  
 comprehensible, 9, 333, 335, 348  
 Hypothesis, 9, 228, 355, 358, 411, 440, 452  
 modified, 384, 385, 393, 397, 465, 466
- passé composé, 416–419, 479, 491, 505
- “passive (construction; formation)”, 16, 18, 153, 159, 235, 244–247, 254–261, 353, 354, 359–378
- pause, 32–39, 44, 113, 387, 428, 464, 489
- Planning, 3, 8, 12, 13, 16, 31–36, 39, 41, 47, 235, 247, 248, 261, 262, 282, 322, 451, 452
- Possessive determiners, 13, 15, 16, 99, 109–124, 199, 201–222
- present perfect, 415–426
- problem-solving task, 8, 282, 358, 414, 120, 485
- Proceduralization, 8, 34, 35, 103, 106, 120, 415
- Question (formation/forms), 13–15, 110, 111, 124, 166, 199, 202, 206, 208–210, 224, 227, 229, 295–301
- Rate of acquisition/learning, 175, 287, 313
- recast, 333, 337–339, 341, 345–348
- Reformulation, 17, 114–120, 337, 341, 413, 420, 463
- Relative clauses, 15, 167–171, 175–194, 442–450
- relativization, 15, 20, 167–193, 433, 442, 443, 450
- repair, 13, 19, 34, 36, 99, 106, 107, 335, 342–348, 383, 388–391, 397–400, 411, 412, 416, 420
- Route of acquisition/learning/development, 4, 5, 11
- Rule, 52, 74, 77, 101, 102, 134–143, 154–158, 236–241  
 grammar, 14, 20, 118, 119, 133, 414  
 metalinguistic, 9, 11, 238  
 pedagogical, 14, 156, 158, 193, 238, 239, 243, 263  
 production, 414–417, 421, 424, 427
- Russian, 14, 54, 76, 91, 131–159, 169, 438, 537–541

- Scaffolding, 395, 409, 459  
 sequence of acquisition/learning/development, 4, 5, 22, 65, 87, 110, 167, 204–206, 280, 281, 427, 460, 483, 500, 541  
 simple past, 415–426  
 Single mechanism/system model, 14, 131–135, 155, 156  
 stages of acquisition/learning/development, 4–6, 13, 54, 58, 62, 65, 70, 77, 86, 87, 93, 94, 102, 107, 111–113, 205, 206, 216, 221, 415, 498  
 Study/Year Abroad, 21, 32–34, 45, 106, 495–524, 536, 547, 552  
 swearwords, 22, 531–537, 542–552
- Task-based instruction/learning/  
 approach/syllabus, 272, 277, 280, 300, 408
- teacher discourse, 457–483  
 tense, 19, 21, 131–159, 207, 227, 242, 280, 285, 393, 414–427, 510, 524  
 Universal Grammar, 25, 169, 427, 440, 442, 450  
 uptake, 18, 333, 336, 340–349, 368
- Variation, 106, 465, 467, 495–524, 533, 538  
 Verb, 12, 14, 32–40, 60, 117, 131–146, 153–159, 179, 206–211, 226, 229, 239–244, 291, 384, 414, 418, 420–426, 501, 505, 508–518, 521  
 Vocabulary, 6, 18, 104, 109, 110, 120, 146, 153, 159, 200, 225, 230, 246, 282, 311–325, 408, 536, 550  
 acquisition/learning, 17, 25, 311–322, 483  
 teaching/instruction, 312, 322, 324, 325





## List of Contributors

**Ahlem Ammar**

McGill University  
Canada  
ahlem\_ammam@hotmail.com

**Tatiana Chernigovskaya**

St. Petersburg State University  
Russia  
tatjana@tc3839.spb.edu

**Jean-Marc Dewaele**

Birkbeck College, University of London  
United Kingdom  
j.dewaele@french.bbk.ac.uk

**María del Pilar García Mayo**

Universidad del País Vasco  
Spain  
figmap@lg.ehu.es

**Kira Gor**

University of Maryland, College Park  
USA  
kiragor@mail.umd.edu

**Peter Griggs**

Institut Universitaire de Formation des  
Maîtres de l'Académie de Lyon, France  
griggs@lyon.iufm.fr

**Alex Housen**

Vrije Universiteit Brussel  
Belgium  
Alex.Housen@vub.ac.be

**Martin Howard**

University College Cork  
Ireland  
mhoward@french.ucc.ie

**Heini-Marja Järvinen**

University of Turku  
Finland  
hemajar@utu.fi

**Tsuyoshi Kida**

Université de Provence  
Aix-en-Provence, France  
tsuyoski.kida@lpl.univ-aix.fr

**Folkert Kuiken**

University of Amsterdam  
The Netherlands  
f.kuiken@uva.nl

**Batia Laufer**

University of Haifa  
Israel  
batialau@research.haifa.ac.il

**Patsy M. Lightbown**

Concordia University  
Canada / USA  
patsy.lightbown@verizon.net

**Katja Lochtman**

Vrije Universiteit Brussel  
Belgium  
klochtma@vub.ac.be

**Barbara Menzel**

University of Applied Sciences  
Kaiserslautern, Germany  
b.menzel@verw-kl.fh-kl.de

**Michel Pierrard**

Vrije Universiteit Brussel  
Belgium  
Michel.Pierrard@vub.ac.be

**Leila Ranta**

University of Alberta  
Canada  
leila.ranta@ualberta.ca

**Ronald Sheen**

University of Quebec in Trois Rivières  
Canada  
rsheen@utcr.ca

American University of Sharjah  
United Arab Emirates  
rsheen@aus.ac.ae

**Nina Spada**

University of Toronto  
Canada  
nspada@oise.utoronto.ca

**Liz Temple**

University of New South Wales  
Australia  
melusine@ihug.com.au

**Siska Van Daele**

Vrije Universiteit Brussel  
Belgium  
Siska.Van.Daele@vub.ac.be

**Ineke Vedder**

University of Amsterdam  
The Netherlands  
i.vedder@uva.nl

**Joanna White**

Concordia University  
Canada  
jwhite@education.concordia.ca



