Developmental Psycholinguistics

On-line methods in children's language processing

Edited by
Irina A. Sekerina
Eva M. Fernández
Harald Clahsen

Language Acquisition & Language Disorders
44

John Benjamins Publishing Company

Developmental Psycholinguistics

Language Acquisition and Language Disorders (LALD)

Volumes in this series provide a forum for research contributing to theories of language acquisition (first and second, child and adult), language learnability, language attrition and language disorders.

Volume 44

Developmental Psycholinguistics. On-line methods in children's language processing
Edited by Irina A. Sekerina, Eva M. Fernández and Harald Clahsen

Developmental Psycholinguistics

On-line methods in children's language processing

Edited by

Irina A. Sekerina
City University New York

Eva M. Fernández
City University New York

Harald Clahsen
University of Essex

John Benjamins Publishing Company
Amsterdam / Philadelphia

Library of Congress Cataloging-in-Publication Data

Developmental psycholinguistics : on-line methods in children's language processing /
edited by Irina A. Sekerina, Eva M. Fernández, Harald Clahsen.
p. cm. (Language Acquisition and Language Disorders, ISSN 0925-0123 ; v. 44)
Includes bibliographical references and index.
1. Language acquisition--Data processing. 2. Language acquisition--Research--Methodology. I. Sekerina, I. A. (Irina A.), 1961- II. Fernández, Eva M. III. Clahsen, Harald.
P118.3.D487 2008
401'.930285--dc22 2007038990
ISBN 978 90 272 5304 0 (Hb; alk. paper)
ISBN 978 90 272 5305 7 (Pb; alk. paper)

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands
John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

Table of contents

Introduction

Irina A. Sekerina, Eva M. Fernández and Harald Clahsen

The study of child language occupies a unique place in research on children's cognitive development. This comes as no surprise, as language is quite close to the core of what it means to be human. Children successfully learn their native language in a relatively short time and without the need for formal instruction. Language is also the main vehicle by which we learn about other people's thoughts; therefore, cognitive and linguistic aspects of human development must be intimately related.

Traditional methods of inquiry in modern linguistics and cognitive psychology have enabled us to learn a great deal about how children acquire language and the stages they go through on their way to adult competence (Pinker 1995). But empirical studies on how children's language develops published over the last 30 or 40 years have a striking characteristic in common: they treat language acquisition as a process that involves building a static database called *the grammar*, to the exclusion of the mechanisms that operate in real time when the child produces or comprehends language. The classic Competence/Performance distinction (e.g., Chomsky 1964) provides a useful framework for discussing this problem: while investigations of child language acquisition are grounded on the assumption that knowledge of language is put to work via a set of processing mechanisms (performance), the primary concern in acquisition research has been with how that principled knowledge (competence) develops. McDaniel, McKee and Cairns (1996), in their seminal book on assessing child language, described how the knowledge that constitutes competence had up until then been extensively studied, and they documented the predominance of *off-line* experimental methods, that is, techniques that prompted children to act out sentences, answer questions or provide grammaticality judgments, responses that could then be compared to those provided by adults or by older or younger children. Armed with empirical evidence of that sort, the field was able to begin to address some of the most basic questions about language development and to formulate explicit descriptions about the nature of developmental sequences.

The era of traditional research on language acquisition, captured so well in the volume by McDaniel et al. (1996), has grown into a mature area of inquiry whose insights have led to a rich understanding about the development of linguistic competence. But times have changed, as we enter a new era that takes a "dynamic processing approach" to the study of language development (Trueswell this volume). We are witnessing a growing interest in the mechanisms that underlie production and comprehension abilities in children, a shift from a focus on competence to a focus on performance. This enterprise has been significantly facilitated by recent advances in technologies that permit tracking behavior at a very fine temporal resolution, methods that have been successfully and extensively applied to study language processing in adults. Such new techniques, which we will collectively refer to as *on-line*, measure reaction times, track eye gazes, examine brain activity. Some of these methods, like self-paced reading, self-paced listening, and cross-modal priming benefit from having a long-standing tradition in the study of adult language processing. Others, like eyetracking and neurophysiological techniques (Henderson & Ferreira 2004; Trueswell & Tanenhaus 2005; Carreiras & Clifton 2004), are newer but quite powerful additions to the experimental toolkit, particularly because they provide the means to study in great detail very early phases of processing, and because they rely little on conscious attention to or metalinguistic awareness of linguistic stimuli.

On-line methods have made their way into language acquisition research with a truly amazing speed. A mere 10 years ago, as documented by Cecile McKee (1996) in her chapter on on-line methods in child language research, reaction time methods (cross-modal priming in particular) dominated the scene, neuroimaging hardly having a presence. Eyetracking was fully absent from McKee's chapter.

When applied to the study of child language, on-line methods permit researchers to observe the interaction of grammar principles (competence) and behavioral limitations and/or preferences (performance), with a greater level of detail and a greater number of perspectives than ever before. We can now investigate how children coordinate multiple sources of information in real time and arrive at sentence meaning using information extracted not only from the words and structure of the sentence but also from the nonlinguistic context. The application of on-line methods also makes it possible to test children's performance limits, to separate performance from competence in assessing children's static and developing linguistic knowledge, an approach that permits building and testing theories about how children's language processing contributes to their acquisition of language (Fodor 1998).

The growing importance of on-line methods in child language research was evident at the forum that brought this volume into being, the *Workshop on On-*

Line Methods in Children's Language Processing held at the Graduate Center of the City University of New York in March 2006. Workshop participants discussed and evaluated questions about the design, methodology, ethics, and practicalities of conducting such studies with children, and speculated on future directions for the emerging field of *developmental psycholinguistics* (Trueswell this volume) and its subfield, *developmental cognitive neuroscience* (Männel & Friederici this volume). In assembling this volume, we asked distinguished researchers – pioneers in the application to child language research of a range of on-line methods – to provide overviews on how changing research paradigms are advancing our understanding of language processing in children. While the overarching theme of this volume is methodological in nature, the collection of chapters achieves a broad coverage also of linguistic and developmental areas by including research on both comprehension and production; by addressing sound-, word- and sentence-level representations; and by discussing aspects of acquisition throughout the entire span of early childhood, from infancy to the elementary school years. The chapters in the volume are dedicated to reaction time methods (Clahsen); eyetracking in its two main forms, free-viewing (Trueswell; Snedeker & Thothathiri) and looking-while-listening (Fernald, Zangl, Portillo & Marchman); and event-related potentials (ERPs; Männel & Friederici). Functional neuroimaging (fMRI), magnetoencephalography (MEG) and optical imaging have yet to make their way into developmental psycholinguistics and, therefore, are not represented in the volume.

We have chosen to group and order the chapters in terms of the methods they focus on, starting with methods examining behavioral responses and followed by methods analyzing event-related potentials and methods tracking eye gazes. Closing the volume, Chapter 6 provides a historical backdrop and speculates about the future of the field.

Chapter 1, "Behavioral methods for investigating morphological and syntactic processing in children" (Harald Clahsen), describes and evaluates experiments using response-time measures to examine processes involved in children's processing of sentences and inflected words. The chapter builds on Cecile McKee's documentation (1998) of on-line methods in child language research and presents an updated overview focusing on techniques that Clahsen, Felser, and the research group at the University of Essex have used to examine how children process complex syntactic phenomena and morphologically complex words in real time. The chapter introduces five criteria against which the various methods for studying children's on-line language processing can be assessed. These criteria are: (a) time sensitivity of the technique, (b) naturalness of stimuli presentation, (c) child appropriateness of the technique, (d) linguistic versatility, and (e) filed compatibility. It then provides an overview of behavioral tasks for investigating children's grammatical processing in production and comprehension. Three experimental techniques are presented in detail: *speeded production*, used to examine automatic processes involved in the spoken production of inflected words; *self-paced listening*, employed in examining children's processing of temporarily ambiguous sentences, specifically relative-clause attachment preferences; and *cross-modal priming* used to determine whether dislocated constituents (e.g., fronted *wh*-phrases) are reactivated at corresponding gap positions during processing. Despite the availability of ERPs and eyetracking, behavioral methods such as the ones outlined by Clahsen still have an important place in developmental psycholinguistics, as these techniques can be applied to study a range of complex and interesting language phenomena, providing time-sensitive measures that require minimal technical equipment.

Chapter 2, "Event-related brain potentials as a window to children's language processing: From syllables to sentences" (Claudia Männel and Angela Friederici), provides an overview of the comprehensive research program to study neurocognition of language, cortical networks and cognitive functions, and language acquisition using ERPs at the Institute of Neuropsychology of the Max-Planck Institute for Human Cognitive and Brain Science in Leipzig. Angela Friederici, head of the Institute, is a leading researcher in applying measures of brain activity in adult sentence processing and has proposed an influential neurocognitive model of language and its extension to the area of language acquisition, which she herself calls *the developmental cognitive neuroscience of language* (Friederici 2000, 2002).

In their chapter, Männel and Friederici describe the successful application of ERP methods to study language acquisition in infants from birth to three years. The chapter describes five ERP components closely linked to language, and how they reflect the processing of phonological, semantic, and syntactic information in progressively older children, compared to adults. Männel and Friederici outline ERP research on a number of landmarks of child language acquisition and identify neural correlates for developmental stages in auditory language comprehension. The research findings summarized in Chapter 2 include (a) work on syllable and stress discrimination in infants using the passive oddball paradigm, (b) investigations of the N400 component as a reflex of phonotactic knowledge, early word learning and knowledge of selectional restrictions for verbs, (c) research measuring sensitivity to sentence-level prosodic cues with the Closure Positive Shift (CPS), and (d) studies eliciting an adult-like biphasic ELAN-P600 component in response to phrase structure violations at the sentence level. Männel and Friederici draw the chapter to a close by demonstrating how ERP components can be used to identify infants at risk for later developing language problems, such as Specific Language Impairment (SLI) and dyslexia. The technique therefore constitutes a new diagnostic tool for very early identification of children who would benefit from intervention.

Chapter 3, "Using eye movements as a developmental measure within psycholinguistics" (John Trueswell), is the first of the three chapters dedicated to eyetracking. Trueswell describes and evaluates free-viewing eyetracking – also known as the visual world paradigm and the world-situated eye-gaze paradigm – to study sentence-level comprehension in toddlers and preschool-age children. John Trueswell and his team at the Institute for Research in Cognitive Science at the University of Pennsylvania were the first to adapt this method to investigate how 5-year-old children comprehend syntactically ambiguous sentences and how sentence processing mechanisms develop (Trueswell, Sekerina, Hill & Logrip 1999). Trueswell begins his chapter with a history of eyetracking in adult research and continues with a technical description of head-mounted, remote and "poor man's" eyetrackers; an explanation of calibration procedures; and remarks about eyetracking data analysis. Trueswell then describes three linking assumptions critical for making valid inferences about what children's eye gaze patterns reveal about the development of sentence processing mechanisms, and, in particular, referential processing.

In a second part of his chapter, Trueswell discusses recent findings from early oculomotor development, visual search, and neurocomputation models of visual attention, all of which are informative with respect to understanding characteristics of spatial attention from infancy until the age of 3. Such discussion provides a solid backdrop for Trueswell's review of the experimental evidence accumulated over the past decades about using eye movements to infer how children resolve prepositional phrase attachment ambiguities, pronominal reference, and quantifier scope, as well as the influence of discourse factors in referential communication tasks.

Chapter 4, "How infants look as they listen: Using eye movements to monitor on-line comprehension by very young language learners" (Anne Fernald, Renate Zangl, Ana Luz Portillo and Virginia Marchman), takes on a complementary approach to the foundational chapter by Trueswell, walking the reader through a detailed description of the listening-while-looking paradigm pioneered by Anne Fernald, the leading author of this chapter, and the Stanford University Center for Infant Studies. Listening-while-looking (LWL) is a version of free-viewing eyetracking adapted for use with infants. The technique was developed out of a desire to overcome a number of shortcomings of commonly used off-line methods to examine language comprehension by infants, such as diary studies, parental-report checklists of vocabulary growth, experiments on word learning, and early versions of the preferential-looking paradigm. These methods do not tap into the real-time properties of spoken language and reveal little about the child's developing efficiency in processing continuous speech during comprehension. The LWL method involves videotaping infants' eye movements and coding these at the finest level of resolution possible in relation to relevant points in the speech signal. The technique can be used to collect eye movement data from infants as young as 14 months, comparable in reliability and precision to data from adult studies that require technically much more sophisticated eyetracking technology. The chapter focuses on the LWL paradigm at a functional level and discusses the logic of each step in the procedure, from preparing and running an experiment to coding eye movements and analyzing the data for several different measures of efficiency in spoken word recognition in infants.

Chapter 5, "What lurks beneath: Syntactic priming during language comprehension in preschoolers (and adults)" (Jesse Snedeker and Malathi Thothathiri) – the third and final contribution on eyetracking – describes how to use a "poor man's" free-viewing eyetracker in testing a particular theoretical problem in the area of language acquisition, i.e., the nature of young children's abstract grammatical representations. Snedeker and Thothathiri exploit the phenomenon of structural priming in production, well known in adult psycholinguistics, to study structural priming during spoken language comprehension in 3- and 4-year-old children.

Snedeker and Thothathiri's starting point is a brief summary of three prior experiments that extend to young children basic findings in structural priming in adult production. This line of research suggests that children older than 4.5 show robust evidence for abstract structural representations of both dative and transitive constructions, while 3-year-olds and younger 4-year-olds do not. Snedeker and Thothathiri set out to test the hypothesis that even younger preschoolers do indeed construct abstract structural representations, seeking evidence of this in their eye movement patterns. The remainder of the chapter reports a series of three experiments employing the "poor man's eyetracker" to examine structural priming in sentences with verbs such as *give* and *hand*, which license dative alternation (V NP NP vs. V NP PP). As expected, eye movement patterns for young 4-years-olds and 3-year olds – as for adults and 4-year-olds – show an effect not only of lexical within-verb priming but, critically, also of abstract between-verb priming for dative sentences.

The final chapter of the volume, "Language acquisition research. A peek at the past: A glimpse into the future" (Helen Smith Cairns), reviews a range of landmark studies in language acquisition research, providing a historical perspective that promotes a better understanding of the significance of the shift in focus from competence toward performance, from off-line to on-line paradigms. Cairns tells the story of how the (still young) field of research on language acquisition has evolved, from early fieldwork studies focusing on the regularities of speech produced by children through studies employing sophisticated off-line techniques to probe underlying linguistic knowledge, though studies taking advantage of on-line techniques to understand performance mechanisms in children. The chapter addresses the complexity of the relationship between theory and practice. Theories have driven experimental innovation, while at the same time the availability of experimental techniques promotes the development of new theories.

Cairns discusses the early interest on the problem of how an underlying grammatical system is acquired. The realization that the speech produced by children vastly under-represents what they might know led to a shift in interest on how implicit knowledge of language develops, an undertaking sustained by the increased sophistication of theories of grammar.

Cairns then provides a brief history of how psycholinguistics developed into a ripe area of inquiry exploring the mechanisms employed in the production and perception of language in adults. As for children, Cairns notes that the preoccupation with competence in language acquisition research, along with a concern to control "performance factors" (such as effects of memory or task demands), resulted in a paucity of studies of children's performance.

But, as this volume represents, and Cairns discusses at length in her chapter, new questions are emerging that directly address the nature of children's language processing. Cairns reviews a range of studies – some discussed elsewhere in this volume, some presented at the Workshop, and others sampled from the literature – that are addressing questions about whether children construct representations that resemble those constructed by adults, about how children revise initial parses, about how children's memory spans limit their performance, about how cross-linguistic research is identifying universal tendencies in child language processing, and about how some of these techniques can be employed for the early detection of language disorders.

To conclude her chapter, Cairns offers some speculations about what the future holds for research in child language development, echoing a number of the other contributions to this volume when stressing the need for multiple and complementary methodological approaches. Progress is called for in particular with respect to the question of how children operate on input to create new grammars, on how adult-like processing skills develop, and on the underlying neural organization. We hope that this volume will lead to innovation in these and related questions.

We cannot understate the importance of the source for this volume, the *Workshop on On-Line Methods in Children's Language Processing* held at the Graduate Center of the City University of New York, March 22–23, 2006. The abstract proceedings for the papers and posters presented at the Workshop are presently available at http://www.qc.cuny.edu/~efernand/childlang/. The first joint scientific gathering specifically dedicated to the emerging field of experimental developmental psycholinguistics, the Workshop gathered specialists in language acquisition, psycholinguistics, neuroscience, and speech-language pathology. Ostensibly, the objective was to provide a forum for discussing the advantages and shortcomings of using on-line methods to study language processing in children, ranging from behavioral methods to paradigms involving eyetracking, to neurophysiological techniques. Beyond discussing methodology, the Workshop initiated dialog between an international and interdisciplinary group of scholars, who were afforded the opportunity to reflect on past landmarks of research in language acquisition, summarize the current state of emerging research on language processing in children, and engage in lively debates about future directions.

The current volume offers six papers loosely based on talks delivered at the Workshop. We have asked the authors to concentrate on methodological matters, but they have gone beyond that directive and produced chapters that serve as more than introductions to experimental paradigms, since they address some of the complex theoretical debates as well as provide a solid overview of child language development.

We are happy to take this opportunity to express our gratitude to those who made the Workshop and this volume possible.

First, we thank the 120 presenters and attendees from the United States, Europe, Japan and Australia for their thought-provoking papers and posters.

Second, we thank the co-directors of the Annual CUNY Human Sentence Processing Conference, Janet Dean Fodor and Diane C. Bradley, for enthusiastically supporting this project and sharing the CUNY Conference infrastructure with us.

In organizing the Workshop and editing the volume, we were assisted by a committee of committed reviewers that included Jennifer Arnold, Holly Branigan, Patricia Brooks, Edward Gibson, Jeffrey Lidz, Eric Pakulak, Julie Sedivy, Valery Shafer, Matthew Traxler, and Virginia Valian. We wish to thank them, along with all the contributing authors, for sharing their expertise and keeping up with our tight deadlines.

We would also like to express our gratitude to Lydia White, Co-Editor of the Language Acquisition and Language Disorders series, Kees Vaes and the John Benjamins Publishing Company, for encouragement and assistance at all stages of production.

Finally, are deeply grateful to the National Science Foundation for its generous financial support of this Workshop through its Child Research Initiative grant #0518438. (The opinions expressed in the Workshop are those of the organizers, presenters and attendees and do not necessarily represent the views of the National Science Foundation or any other governmental organization.)

References

Carreiras, M. & Clifton, C., Jr. (Eds.). (2004). *The on-line study of sentence comprehension: Eyetracking, ERPs and beyond.* Hove: Psychology Press.

Chomsky, N. (1964). [The development of grammar in child language]: Discussion. In U. Bellugi & R. Brown (Eds.), *The acquisition of language*, Vol. 1: *Monographs of the Society for Research in Child Development, 29*(1), 35–42.

Fodor, J. D. (1998). Parsing to learn. *Journal of Psycholinguistic Research, 27*(3), 339–374.

Friederici, A. D. (2000). The developmental cognitive neuroscience of language: A new research domain. *Brain and Language, 71,* 65–68.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences, 6*(2), 78–84.

Henderson, J. M. & F. Ferreira (Eds.). (2004). *The interface of language, vision, and action: Eye movements and the visual world*. Hove: Psychology Press.

McDaniel, D., McKee, C., & Cairns, H. S. (Eds.). (1996). *Methods for assessing children's syntax.* Cambridge MA: The MIT Press.

McKee, C. (1996). On-line methods. In D. McDaniel, C. McKee, & H. S. Cairns (Eds.), *Methods for assessing children's syntax* (pp. 189–212). Cambridge, MA: The MIT Press.

Pinker, S. (1995). Language acquisition. In L. R. Gleitman & M. Liberman (Eds.), *An invitation to cognitive science,* Vol. 1.: *Language* (2nd ed., pp. 135–182). Cambridge MA: The MIT Press.

Trueswell, J. C. & Tanenhaus, M. K. (Eds.). (2005). *Approaches to studying world-situated language use: Bridging the language-as-product and language-as-action traditions*. Cambridge MA: The MIT Press.

Trueswell, J. C., Sekerina, I. A., Hill, N., & Logrip. M. (1999). The Kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition, 73*, 89–134.

List of contributors

Helen Smith Cairns
Queens College and Graduate Center
City University of New York
65-30 Kissena Blvd.
Flushing, NY 11367-1597
Tel: 718-997-2870
Fax: 718-997-2873
hcairns@optonline.net

Harald Clahsen
University of Essex
Colchester, CO4 3SQ, UK
Tel: +44/1206/87-2228
Fax: +44/1206/87-2198
harald@essex.ac.uk

Anne Fernald
Stanford University
Stanford, CA 94305
Tel: 650-725-2400
Fax: 650-725-5699
afernald@stanford.edu

Eva M. Fernández
Queens College and Graduate Center
City University of New York
65-30 Kissena Blvd.
Flushing, NY 11367-1597
Tel: 718-997-2867
Fax: 718-997-2873
eva.fernandez@qc.cuny.edu

Angela D. Friederici
Max Planck Institute for Human
Cognitive and Brain Sciences
Stephanstr. 1a
04103 Leipzig, Germany
Tel.: +49 /0341/99-40112
Fax: +49/0341/99-40113
angelafr@cbs.mpg.de

Claudia Männel
Max Planck Institute for Human
Cognitive and Brain Sciences
Stephanstr. 1a
04103 Leipzig, Germany
Tel.: +49 /0341/99-40112
Fax: +49/0341/99-40113
maennel@cbs.mpg.de

Virginia A. Marchman
Stanford University
Stanford, CA 94305
Tel: 650-725-2400
Fax: 650-725-5699
marchman@psych.stanford.edu

Ana Luz Portillo
Stanford University
Stanford, CA 94305
Tel: 650-725-2400
Fax: 650-725-5699
analuz@stanford.edu

Irina A. Sekerina
College of Staten Island and Graduate
Center, City University of New York
2800 Victory Blvd.
Staten Island, NY 10314-6609
Tel: 718-982-3760
Fax: 718-982-4114
sekerina@mail.csi.cuny.edu

Jesse Snedeker
Harvard University
William James Hall, 33 Kirkland Street
Cambridge, MA 02138-2044
Tel: 617-495-3873
Fax: 617-384-7944
snedeker@wjh.harvard.edu

Marathi Thothathiri
Harvard University
William James Hall, 33 Kirkland Street
Cambridge, MA 02138-2044
Tel: 617-384-8357
Fax: 617-384-7944
malathi@wjh.harvard.edu

John C. Trueswell
University of Pennsylvania
3401 Walnut Street, Room 302C
Philadelphia, PA 19104-6228
Tel: 215-898-0911
Fax: 215-573-9247
trueswel@psych.upenn.edu

Renate Zangl
Stanford University
Stanford, CA 94305
Tel: 650-725-2400
Fax: 650-725-5699
ezangl@stanford.edu

CHAPTER 1

Behavioral methods for investigating morphological and syntactic processing in children

Harald Clahsen

While most first language acquisition research to date has focused on the development of children's linguistic competence, a number of research teams have also investigated the mechanisms children employ to process sentence-level and word-level information in real time, by applying experimental techniques familiar from the adult processing literature to children. This chapter presents an overview of different kinds of behavioral tasks for investigating both morphological and syntactic processing in children focusing on three techniques that we have explored in our own research on children's on-line language processing: self-paced listening, cross-modal priming, and speeded production.

1. Introduction

In 1996, Cecile McKee presented an overview of the very small number of on-line techniques suitable for studying syntactic processing in children that were available at the time. The purpose of this chapter is to provide an updated overview. My focus will be on children's grammatical processing and on different kinds of *behavioral* tasks for investigating morphological and syntactic processing in children.

Language processing can be conceived of as a sequence of operations, each of which transforms a linguistic representation of a stimulus into a linguistic representation of a different form. Research into language processing examines how linguistic representations are constructed *in real time* during the comprehension and production of language and how different sources of information become available over time. To study the processes involved in production and comprehension as they occur, time-sensitive, so-called on-line, measures of language processing are required. The advantages of using on-line experimental techniques are that they allow us to tap into automatic unconscious processes involved in language comprehension and production and that they minimize participants' reliance on explicit or metalinguistic knowledge. There are two basic types of time-sensitive measures available to examine language processing: behavioral measures (e.g., comprehension response times and production latencies) and physiological measures (e.g., event-related brain potentials (ERPs) and eye movements). As the latter will be presented in other chapters of this book (see Männel & Friederici for ERPs, and Trueswell for eye movement experiments), I will only be concerned with behavioral measures of language processing here.

Before presenting an overview of behavioral experimental methods, it is necessary to establish some criteria against which the various methods can be evaluated. The first criterion concerns the *time-sensitivity* of a technique and asks at which point in time during language processing a particular measure is taken. Clearly, if a technique measures responses at the offset of a stimulus, e.g., at the end of a sentence, it is not particularly revealing for understanding the moment-by-moment characteristics of processes occurring during the processing of that sentence. The second criterion is whether the stimuli are presented *in a natural way* allowing participants to process them using normal listening or reading. As we will see, this is particularly difficult to achieve for behavioral experimental tasks. The third question we ask is whether the experimental task assigned to participants is *child-appropriate*. Some techniques require a dual task, e.g., monitoring for a visual target while listening to a sentence, which may be too challenging for young children. The fourth criterion is whether a technique is *linguistically versatile*, i.e., applicable to a range of different linguistic phenomena. Finally, we will ask whether a technique is *field-compatible*. This refers to practical considerations in running experiments with children. In some circumstances, for example, it is impossible to bring children into the lab. In such cases, it would be advantageous if a technique required minimal equipment so that children can be tested at their schools or their homes.

In the remainder of this chapter, I will consider behavioral methods first for studying on-line sentence comprehension and second for investigating language production. My focus will be on the advantages and disadvantages of three techniques that we have explored in our own research on children's on-line language processing: (i) the self-paced listening task to examine children's comprehension of ambiguous sentences; (ii) the cross-modal picture priming task to study children's comprehension of syntactic dependencies, specifically *wh*-dependencies, and (iii) the speeded production task to investigate processes involved in children's production of morphologically complex words.

2. Behavioral methods for studying grammatical comprehension

The adult psycholinguistic literature offers a range of behavioral methods for investigating on-line grammatical comprehension, but only a small number of techniques have been used with children: word monitoring during sentence comprehension, probe recognition, speeded grammaticality judgment, self-paced reading and listening, and cross-modal priming. What is common to these techniques is that they can be used with children from about 4 or 5 years of age onwards to study relatively complex syntactic phenomena. The study of language processing in infants requires different techniques measuring, for example, preferential looking and head-turning patterns (see Fernald, Zangl, Portillo, & Marchman this volume). Here, I will first briefly present word monitoring, probe recognition, and speeded grammaticality judgment, and then discuss in some more detail self-paced reading, self-paced listening and cross-modal priming.

2.1 Word monitoring

Tyler and Marslen-Wilson (1981) were among the first to investigate on-line sentence comprehension in children. They used a task in which participants monitor linguistic stimuli, e.g., auditorily presented sentences as such as those in (1), for a particular target word, e.g., the word *hand*. The participant's response, usually either a button press or a vocal response, indicates that the participant has noted the occurrence of the target in the sentence. Word-monitoring times are measured from the target's occurrence in the sentence to the participant's response.

(1) a. John had to go back home. He had fallen out of the swing and had hurt his *hand* on the ground.
b. John had to sit on the shop. He had lived out of the kitchen and had enjoyed his *hand* in the mud
c. The on sit top to had John. He lived had and kitchen the out his of had enjoyed *hand* mud in the

Tyler and Marslen-Wilson (1981) applied this technique to 5-, 7-, and 10-year-old children, and found that the children's ability to detect a word target showed the same pattern of degradation as it did in adults with the shortest monitoring times for contextually appropriate sentences (1a), followed by contextually inappropriate sentences (1b), and semantically and syntactically anomalous sentences (1c). This finding was taken to indicate that children and adults analyze sentences in essentially the same way, i.e., children like adults use context information to construct interpretative representations on-line which in turn facilitates the recognition process of the target words. In the category monitoring task, however, in which children had to monitor the sentences for a member of a particular semantic category (e.g., 'Monitor for body parts'), the 5-year-olds differed from the older children in that the facilitating effect of contextually appropriate sentences (1a) was smaller than for the 7- and 10-year-olds. Tyler and Marslen-Wilson (1981) attributed this finding to the additional processing cost associated with semantic-attribute matching – possibly in conjunction with a more general problem with utilizing certain types of pragmatic cues during sentence comprehension.

Tyler and Marslen-Wilson's (1981) word-monitoring task allows the researcher to examine the role of different kinds of contextual information for word recognition. The task provides a time-sensitive measure of word recognition in context and allows listeners to process the auditorily presented sentences in a normal way. However, as McKee (1996) pointed out, a disadvantage of this technique is that only a limited range of relations between target words and their host material can be studied with this technique. It is, for example, hard to see how coreference relations and other kinds of syntactic dependencies could be examined with this technique.

2.2 Probe recognition

In the probe recognition task, participants hear or read a sentence. At some point, the presentation of the sentence is stopped and participants are asked to decide whether a visually or auditorily presented target word ('probe') had occurred in the preceding linguistic material. Response times are measured from the onset of the probe item to the beginning of the participant's vocal response or button press. Several researchers have used this task to examine different kinds of syntactic dependencies in adult sentence comprehension (Bever & McElree 1988; McElree & Bever 1989; MacDonald 1989; Bever & Sanz 1997). Consider, for example, sentences such as those in (2) from McElree and Bever (1989):

(2) a. The dazed cabbie who drove the beat-up taxi was resented (P1)
constantly (P2).
DAZED

b. The dazed cabbie who drove the beat-up taxi was resentful (P1)
constantly (P2).
DAZED

P1 and P2 indicate the points at which the probe items appeared. McElree and Bever did not find any difference between (2a) and (2b) at P1, but at the end of the sentence (P2), response times were significantly shorter for (2a) than (2b).

Probe recognition times are known to yield faster response times for target words that were recently perceived than for those that are further away from the end of the sentence. Given the assumption that passive sentences such as (2a) contain a syntactic gap of the dislocated object, the shorter probe recognition times to DAZED in (2a) have been interpreted as a recency effect, due to the reactivation of the dislocated phrase *the dazed cabbie* after *resented*.

Mazuka (1998) applied this technique to groups of English and Japanese-speaking children as young as 4. Children had to listen to sentences involving main and subordinate clauses and were probed on auditory word targets from these sentences. Her results indicate differences in the way main and subordinate clauses are processed in the two languages. Specifically, the English-speaking children showed an advantage for subordinate clauses (as revealed by shorter response times in a lexical probe recognition task), whereas the Japanese children had shorter RTs for main clauses.

From a methodological perspective, one crucial disadvantage of the probe-recognition task is that it appears to be less time-sensitive than other on-line techniques and that the task is sensitive to a variety of strategic processes (Gordon, Hendrick, & Foster 2000). In many studies, probe-recognition times are measured at the end of the sentence. These data do not tap on-line syntactic processing as it occurs but are more likely to pick up sentence-final wrap-up processes, which may involve semantic rather than syntactic representations. A disadvantage of studies that measured probe-recognition times at within-sentence test points is that the stimulus sentences have to be interrupted, which makes the task rather unnatural.

2.3 Speeded grammaticality judgment

In this task, participants are asked to judge the grammaticality or ungrammaticality of linguistic stimuli as quickly as possible. Timed or speeded grammaticality judgment tasks have been widely used to examine adults' sensitivity to various types of grammatical and semantic information, or relative processing difficulty. The general assumption is that relative processing difficulty should be reflected in slower response times, lower response accuracy, or both (McElree & Griffith 1995, 1998). A variant of this task, the violation detection paradigm, has also been applied to children (Wulfeck 1993; Kail & Diakogiorgi 1998; Kail 2004). Consider, for example, Kail's (2004) study in which three age groups of French children (mean age: 6;8, 8;6 and 10;10 years) and a group of adults were asked to detect agreement violations in sentences such as (3a) and word order violations in sentences such as (3b):

(3) a. Chaque semaine la voisine *remplissent le frigo
Every week the neighbor[sg.] fill [pl.] the fridge
après avoir fait les courses au marché.
after having done the shopping at the market

b. Chaque semaine *remplit la voisine le frigo
every week fill [sg] the neighbor the fridge
après avoir fait les courses au marché.
after having done the shopping at the market

The stimulus sentences were presented auditorily with normal intonation. Participants were asked to decide whether a sentence had 'good grammar' and were specifically instructed to press a button as soon as they discovered an ungrammaticality. While response times to the grammatical sentences were not analyzed, the response latencies for the ungrammatical sentences were measured from the offset of the word (e.g., **remplissent* in (3a)) that made a sentence ungrammatical. Kail (2004) found that both children and adults were faster in detecting agreement violations than word order violations suggesting differences in sensitivity to different types of ungrammaticality.

From a methodological perspective, it is noteworthy that children's response times in this task were substantially longer than those of adults. This was the case not only for the 8-to-9-year-olds, who had an overall response time of 2017 ms, but also for the 6-to-8-year-olds who had a mean overall response time of 2573 ms, more than three times of the adult group. These extremely long response times suggest that this task is particularly difficult for children and, more importantly, that the data are unlikely to tap automatic processes involved in children's language processing. Moreover, grammaticality judgment tasks have been subject to much criticism as the degree to which such judgments reflect implicit grammatical competence is unclear (Schütze 1996). It is also not obvious how young children interpret the instruction to decide between sentences that have good vs. bad grammar.

2.4 Self-paced reading and self-paced listening

In this task, sentences are presented segment-by-segment or word-by-word either visually or auditorily. Participants trigger the presentation of subsequent segments by pressing a pacing button. In self-paced reading, prior segments or words may either stay on the screen or disappear upon pressing the pacing button. Times between button presses are recorded and provide a step-by-step record of the parse as it unfolds. The basic rationale underlying this task is that increased reading or listening times to a particular segment (relative to the same segment in a control condition) indicate relatively higher processing difficulty at this point in the sentence (Just, Carpenter, & Wooley 1982; Mitchell 2004).

The self-paced reading task has been widely used in adult sentence processing research to investigate a range of phenomena, e.g., the on-line interpretation of temporarily ambiguous sentences (see, e.g., Gibson, Pearlmutter, Canseco-Gonzalez, & Hickock 1996) and different kinds of syntactic dependencies (e.g., Clifton & Frazier 1989), the processing of multi-clausal structures (Gibson & Warren 2004), etc. Results from these studies have shown that adults are guided by different types of information during parsing including phrase-structure information, lexical-semantic information, and contextual information (Mitchell 2004).

There are a few studies that used self-paced reading or listening to examine on-line sentence processing in children. Traxler (2002) studied subject-object ambiguities in English-speaking 8-12-year-olds using the self-paced reading task. The materials included (i) sentences such as (4a) which are known to produce a clear garden-path effect in adults (because *the girl* is initially interpreted as the direct object of *tripped*, an analysis that has to be revised later in the clause), (ii) sentences such as (4b) in which the postverbal NP is a semantically implausible object of the verb, and (iii) sentences such as (4c) that contained intransitive verbs. The control conditions for all cases were corresponding sentences in which the embedded verb and the postverbal NP were separated by a comma, thereby precluding the subject-object ambiguity.

(4) a. When Sue tripped *the girl* <u>fell over</u> and the vase was broken.
b. When Sue tripped *the table* <u>fell over</u> and the vase was broken.
c. When Sue fell *the policeman* <u>stopped</u> and helped her up.

The children's reading times were found to be shorter in the ambiguous region (shown in italics) and longer in the disambiguating region (underlined) relative to the control condition with commas. Like adults (Traxler 2005), 8- to 12-year-old children tended to misanalyze the postverbal NP in all three conditions as a direct object indicating that children (and adults) prefer the structurally simpler analysis irrespective of semantic plausibility. The effect, however, was less strong in the intransitive condition, suggesting that subcategorization information was at least partially utilized.

Sentence processing in pre-literate children can be studied using the self-paced listening technique, in which participants listen to sentences by pressing a pacing button to receive successive words or phrasal segments. This technique has been used successfully with adults (Ferreira, Henderson, Anes, Weeks, & McFarlane 1996; Ferreira, Anes, & Horine 1996; among others) and has been shown to be sensitive to the same effects that have been observed in corresponding tasks using visual stimuli. Booth, MacWhinney and Harasaki (2000) investigated 8- to

12-year-old children's on-line comprehension of relative clauses using both self-paced reading and self-paced listening tasks. Their materials included different kinds of relative clause structures:

(5) a. The monkey that followed the frog left the tree in a hurry.
b. The deer that the tiger watched entered the field from the side.

In (5a), both the antecedent NP and the relativized NP fulfill the grammatical function of subject whereas (5b) contains an object relative. The results revealed a slow-down in both reading and listening times at the relative clause – main clause transition for object relatives (e.g., (5b)) compared to subject relatives (5a) indicating increased processing difficulty for the former. Booth et al.'s findings confirm that self-paced listening and reading yield similar experimental effects, not only in adults, but also in children. Felser, Marinis and Clahsen (2003) and Kidd and Bavin (2007) used the self-paced listening task to investigate how children process ambiguous sentences. Here we will consider the Felser et al. study as an example.

2.4.1 *Investigating relative clause attachment with self-paced listening*

Felser et al. (2003) investigated relative-clause attachment preferences in 6-to-7-year-old children in sentences such as *The doctor recognized the nurse of (with) the pupil who was feeling very tired.* In such sentences, the relative clause can either be interpreted to modify the second noun phrase (NP2 disambiguation) implying that *the pupil* was feeling very tired, the option typically preferred by native speakers of English, or the first one (NP1 disambiguation) implying that *the nurse* was feeling very tired. Previous research on adult native speakers has shown that disambiguation preferences are affected by the type of preposition joining the two potential antecedent noun phrases. NP2 disambiguation is preferred cross-linguistically if the two possible antecedent NPs are joined by a thematic preposition such as *with* (Gilboy, Sopena, Clifton, & Frazier 1995; De Vincenzi & Job 1993; Traxler, Pickering, & Clifton 1998). For antecedent NPs joined by the case-assigning preposition *of* or its translation equivalents, on the other hand, attachment preferences have been found to vary across languages (Carreiras & Clifton 1993, 1999; Cuetos, Mitchell, & Corley 1996; Fernández 2003). One explanation for the robust NP2 preference for NPs linked by semantically contentful prepositions is that prepositions such as *with* create a local thematic domain of their own, and that the parser prefers to associate ambiguous modifiers with material inside local thematic domains (Frazier & Clifton 1996). In the absence of such lexical biases, attachment preferences are determined by other factors including phrase-structure based locality principles such as 'Predicate Proximity' or 'Recency' (Gibson et al. 1996). According to former, ambiguous modifiers are attached as close as possible to the main predicate, yielding NP1 attachment, whereas according to the latter, ambiguous modifiers are attached to the most recently processed constituent, yielding NP2 attachment.

Felser et al. (2003) tested 6- to 7-year-old children and adult controls in a self-paced listening task. The experiment had a 2×2 design with the factors 'Preposition' and 'Attachment' yielding four conditions as illustrated in (6). All experimental and filler sentences were split up into five segments as shown in (7). Disambiguation using grammatical number always occurred on the fourth segment, i.e., on the auxiliary. To ensure that the experimental sentences sounded equally natural in both the *of* and the *with* conditions, the relative ordering of NP1 and NP2 was reversed in the *with* conditions. Additional off-line and on-line control experiments revealed that NP order by itself did not influence attachment decisions.

(6) The doctor recognized…
  a. Of-NP1: …the nurse of the pupils who was feeling very tired.
  b. Of-NP2: …the nurse of the pupils who were feeling very tired.
  c. With-NP1: …the pupils with the nurse who were feeling very tired.
  d. With-NP2: …the pupils with the nurse who was feeling very tired.

(7) The doctor recognized / the nurse of the pupils / who / was / feeling very tired.

After listening to each segment, the participants were asked to press a button on a dual push-button box as quickly as possible in order to receive the next segment. The end of each sentence was indicated by a tone. To ensure that the participants paid attention to the task, all experimental sentences and half of the fillers were followed by a comprehension question, which was also presented auditorily.

Table 1 provides an overview of the different participant groups' mean reaction times to the disambiguating auxiliary.

Only the adult group showed a significant interaction between Preposition and Attachment, indicating that their attachment preferences were influenced by the type of preposition involved. The children differed from the adult controls in that their disambiguation preferences were not affected by the type of prepo-

Table 1. Mean reaction times (in ms) for segment 4 (adapted from Felser et al. 2003).

	Adults	**High-Span Children**	**Low-Span Children**
of-NP1	610	863	874
of-NP2	665	964	807
with-NP1	667	749	916
with-NP2	609	807	836

sition (*of* vs. *with*) at all. Instead, the children's on-line attachment preferences were found to interact with their working memory span. While the high-span children showed a preference for NP1 attachment irrespective of the preposition involved, the low-span children showed an overall preference for NP2 disambiguation. These results are in contrast to the findings from a recent reading study with adults (Swets, Desmet, Hambrick, & Ferreira 2007) in which high-span adults were found to favour local (NP2) attachment of RCs, whereas low-span adults favoured non-local (NP1) attachment. It is not clear whether these discrepancies are due to the different modalities (reading vs. listening) tested or due to differences between children and adults. In any case, Felser et al. argued that during listening, the children applied one of two different phrase-structure based locality principles, depending on their working-memory span. Whereas high-span children follow a 'Predicate Proximity' strategy, low-span children tend to associate the relative clause with the most recently processed NP. Thus, similarly to what Traxler (2002) found, children seem to apply the same kind of phrase-structure based parsing heuristics as adults but are more limited in their ability to exploit lexical-semantic information during on-line ambiguity resolution.

2.4.2 *Methodological issues*

The design of materials for self-paced listening experiments requires particular attention to the prosodic properties of the stimuli and potential intonational cues. One way of addressing this concern is by splicing in the relevant segments from other sentences in order to neutralize as much as possible any intonation biases and to ensure the critical items are acoustically identical in all sentences. For the materials used by Felser et al. (2003), for example, the initial NP of each NP complex was replaced by the same NP taken from another sentence read separately. Additionally, the words *who*, *was*, and *were* were spliced out and replaced by the same words read separately; see Felser et al. (2003: 151ff.) for further discussion of the role of prosody in self-paced listening. Moreover, various pretests are required to control for different factors potentially affecting the results of the main experiment, a vocabulary test to assess whether the children know the vocabulary items that are used in the main experiment, an auditory off-line questionnaire and/or a grammaticality judgment task to ensure that the children are able to comprehend the kinds of sentences used in the main experiment, and, given the effects of working memory seen in this kind of task, a listening-span test to assess children's working memory.

In sum, self-paced reading and listening are useful techniques to examine children's on-line sentence processing. The technique provides a time-sensitive measure, i.e., a segment-by-segment or word-by-word record of sentence processing time. The advantages of self-paced reading and listening are that this technique can be applied to a wide range of linguistic phenomena and that it requires minimal technical equipment (essentially a PC or Laptop and a push-button box), which makes it suitable for use outside the experimental laboratory. For the Felser et al. (2003) study, for example, it was not possible to bring children into the laboratory (as the university was unwilling to cover the required insurance). We therefore had to run the experiments at the children's schools, which could easily be done for a self-paced listening experiment. The task assigned to participants is not particularly demanding, even though we saw some effects of working memory in the children, which might reflect task demands that differ between children and adults. A potential disadvantage of self-paced reading and listening is the segment-by-segment or word-by-word stimulus presentation, which yields relatively slow response times in comparison to, for example, eye movement or ERP experiments and does not allow participants to read or listen to the sentences in the usual way, even though new technologies such as instant messaging, online chats, e-books, podcasts and webcasts make both self-paced reading and self-paced listening more commonplace.

2.5 Cross-modal priming

In this task, participants are required to name or, more commonly, make a lexical decision to visual targets while listening to stimulus words or sentences spoken at normal speed. The rationale is that the processing of visual targets is facilitated if they are presented immediately after the auditory presentation of an identical or semantically related word, or 'prime'. In sentence-processing research, both on adults and children, cross-modal priming has been used to examine the processing of sentence-internal referential dependencies, e.g., binding principles (Nicol & Swinney 1989; McKee, Nicol, & McDaniel 1993), and of filler-gap dependencies such as those in sentences involving *wh*-movement (e.g., Love & Swinney 1996, 2007; Hestvik, Schwartz, Tornyova, & Datta 2005; Roberts, Marinis, Felser, & Clahsen 2007) and object scrambling (e.g., Clahsen & Featherston 1999; Nakano, Felser, & Clahsen 2002).

With respect to binding principles, it has been found that in sentences such as (8) both adults (Nicol & Swinney 1989) and preschool children (McKee et al. 1993) responded faster to visual targets such as LEOPARD in the reflexive than in the non-reflexive condition.

(8) The alligator knows that the leopard with green eyes is patting
himself/him on the head with a pillow.
↑
[LEOPARD]

This contrast suggests that a binding principle (according to which a reflexive pronoun must be bound by a local antecedent within the same clause) affects on-line sentence processing in that coreference between the reflexive and its antecedent is immediately established; see McKee (1996: 195ff.) for a detailed description of the child version of this experiment.

Several studies using cross-modal priming have examined the processing of filler-gap dependencies in adults. Love and Swinney (1996) studied English sentences containing object-relative clauses, such as *Jimmy used the new pen that his mother-in-law recently purchased*, in which the object (*the new pen*) is dislocated from the subcategorizing verb (*purchased*). Love and Swinney (1996) found that lexical decision times on targets appearing at the offset of *purchased*, where the gap is, were significantly shorter for targets that were semantically related to the object of the embedded verb than for unrelated ones, whereas at a control position preceding the verb *purchased*, there was no such difference. These findings indicate that the parser recovers or reactivates the grammatical and semantic features of the dislocated constituent (*the new pen*) at a potential gap site yielding a semantic priming effect at the gap position but not at the control position. An alternative interpretation of these findings is the so-called direct association account according to which a displaced argument will be linked to its subcategorizing verb once this is encountered (Pickering 1993; Traxler & Pickering 1996) yielding antecedent reactivation on or immediately after the main lexical verb (*purchased* in the example above).

Antecedent priming in (4- to 6-year-old) children has been studied by Love and Swinney (2007), Hestvik et al. (2005) and Roberts et al. (2007). Love and Swinney and Roberts et al. adopted the cross-modal picture priming task from McKee et al. (1993) and Hestvik et al. (2005) used an alternative picture-*naming* version. Let us consider the Roberts et al. study more closely.

2.5.1 *Investigating antecedent reactivation in children's sentence processing*

Roberts et al. (2007) examined indirect object-relative clauses such as (9) in groups of 5- to 7-year-old children and adults. These kinds of sentences made it possible (unlike the direct object relative clauses used by Love & Swinney 2007) to test potential antecedent priming effects at gap sites that are not adjacent to the subcategorizing verb:

(9) John saw [the peacock]$_i$ to which the small penguin gave the nice birthday present t_i in the garden last weekend.

For each experimental sentence, there were two visual targets, a 'related target', i.e., a picture of the indirect object noun (e.g., a picture of a peacock for (9)), and an ‘unrelated target’ (e.g., a picture of a carrot for (9)). Visual targets were shown at two positions, (i) at the gap position, i.e., at the offset of the final word of the direct object NP, e.g., after *present* in (9), and (ii) at a control position 500 ms earlier. Each experimental sentence was presented identically to four groups of subjects: the first one saw the related target at the gap position, the second group at the control position; the third group saw the unrelated target at the gap position, and the fourth group at the control position. During the presentation of the sentences, pictures appeared on the computer screen, and the participants were required to decide whether the animal/object in the picture was *alive* or *not alive*, by pushing buttons on a push-button box. Response times were measured from the point at which the picture appeared on the screen to the participant’s pressing of the response button. To ensure that the participants paid attention to the task, they were also asked to respond to (yes-no) comprehension questions randomly interspersed throughout the experiment asking for one of the main characters.

The results from this study showed that the children’s reaction times were slower overall than the adults’ and that children’s and adults’ processing of filler-gap dependencies was affected by working memory differences. For children and adults with a high working memory span, a Position × Target Type interaction was found indicating antecedent reactivation at the gap position in these participants. All high-span participants responded more quickly to identical than to unrelated picture targets at the gap position, and lexical decision times for ‘identical’ were shorter at the gap position than at the earlier control position; these contrasts are illustrated in Table 2.

Low-span children and adults, on the other hand, did not show any antecedent reactivation at the gap position. Interestingly, this did not compromise their ability to understand the experimental sentences, since they answered the comprehension questions that were asked after the auditory stimuli as accurately as the high-span participants. In any case, the finding that working memory is a relevant variable for discovering antecedent reactivation effects is consistent with the results of earlier studies showing that the processing of complex sentences in

Table 2. High-span children and adults’ mean reaction times (in ms) to picture targets (adapted from Roberts et al. 2007).

	High-Span Children (N = 19)		**High-Span Adults (N = 22)**	
	Control Position	**Trace Position**	**Control Position**	**Trace Position**
Identical Targets	1245	1158	694	678
Unrelated Targets	1158	1211	692	709

general, and of filler-gap constructions in particular, incurs a working memory cost in adults (Gibson & Warren 2004; King & Kutas 1995; Miyamoto & Takahashi 2001; Nakano et al. 2002) and that for children, memory capacity may be a predictor of effective language processing (e.g., Booth et al. 1999, 2000; Gathercole & Baddeley 1989). One consequence of this is that studies of sentence processing in children (particularly of complex sentences) should be accompanied by a working memory test (along with other pretests).

2.5.2 *Methodological issues*

Cross-modal priming studies and the conclusions drawn from these studies have been subject to methodological criticism in the past. Specifically, McKoon, Ratcliff and Allbritton (1996) and McKoon and Ratcliff (1994) present experiments suggesting that apparent antecedent reactivation effects may be artifacts of the particular method used for selecting control words. That is, the reason why semantically related (as opposed to unrelated) targets often trigger shorter reaction times may simply be that they fit better into the current sentential context, and hence can be integrated more easily than poorly-fitting control words. Note, however, that the evidence for a 'goodness-of-fit' effect presented by McKoon and Ratcliff (1994) comes from a different task, unimodal instead of cross-modal presentation, and that even though McKoon et al. (1996) used cross-modal priming, the presentation rate of the spoken sentences was extremely slow, 390 ms per word. Nicol, Swinney, Love and Hald (2006) replicated the goodness-of-fit effects for a unimodal presentation paradigm as well as for a slowed speed cross-modal task, but not for the commonly used cross-modal priming task that uses continuous sentence presentation. Nicol et al. (2006) therefore concluded that 'goodness of fit' does not influence lexical decision times in the cross-modal priming task and that this technique is indeed sensitive to on-line syntactic parsing rather than to artificial integration processes.

Another methodological issue concerns the way the visual targets are related to the primes in the auditory stimulus material. In most studies using cross-modal priming, the experimental target words are usually strong semantic associates of the antecedent, whereas the control targets are semantically unrelated to the antecedent. One problem with this is that on top of antecedent reactivation, an additional processing step is required to establish a semantic association between the syntactic gap and the target word (see Clahsen & Featherston 1999). An alternative is to use identical repetitions as visual targets, i.e., the same word as the dislocated antecedent, e.g., *peacock* in (9). Given that the 'gap' can be conceived of as containing a silent copy of the displaced constituent, using identical targets is the most direct way of testing whether or not such a copy forms part of the mental representation of the sentence during on-line processing. The main disadvantage of using the actual antecedent is that participants might realize that targets were preceding words and start to anticipate this, which would change the nature of the task, making it a conscious recall task rather than an unconscious measure of on-line processing. This can be avoided, however, by using a larger proportion of unrelated targets than usual. In studies that use identity targets, the proportion of sentences with identical targets to those with unrelated targets is about one in twenty, making conscious detection of repeated words an unlikely possibility. Moreover, any amount of priming due to the formal or semantic identity of the antecedent and the target can be factored out by comparing priming effects at the gap position with those on the same target word at control positions.

Detailed methodological advice for constructing a cross-modal priming experiment is given in McKee (1996). In addition to the points mentioned there, a number of pretests are required for the construction of appropriate materials and to rule out potentially confounding factors. For a picture-priming experiment, a picture-classification task is necessary to ensure that the children are able to correctly classify the target pictures as 'alive' or 'not alive'. Moreover, if complex sentences such as those in (9) are to be examined in the main experiment, the children's ability to comprehend these kinds of sentences needs to be pre-tested along with their working memory span.

Cross-modal priming offers some advantages over other behavioral methods for studying on-line sentence processing in children. It allows for the stimulus materials to be presented uninterrupted and at a normal speech rate, thus rendering it more natural than, for example, self-paced reading or listening. The use of picture targets instead of written words makes the technique suitable for young children who cannot yet read or write. Moreover, the possibility of presenting visual targets at different positions during the auditory stimulus allows for the discovery of potential priming effects occurring on-line at specific positions in the sentence. Another advantage is that participants' attention is directed towards the decision task on the visual targets and away from the sentence stimuli that the researcher is interested in, thereby reducing the possibility of participants' responses being influenced by conscious processes.

One disadvantage of this technique is that, unlike ERPs or eye movement measures, cross-modal priming does not provide a continuous record of sentence processing, but instead more of a snapshot view of the state of the language processor at particular positions in the sentence, namely at the specific test points at which visual targets are presented. Another disadvantage of cross-modal priming is the complexity of the task, attention to the auditory sentence while performing a decision task on a picture target. These dual-task demands might have been the reason why Roberts et al. (2007) failed to detect antecedent reactivation in adults and children with low working memory. As an alternative, Hestvik et al. (2005)

have proposed a supposedly more child-friendly picture-naming version of the task in which pictures shown during listening to the sentences have to be named. Hestvik et al. found antecedent reactivation effects in 8- to 11-year-old children, not only for those with high but also for those with low working memory. Note, however, that their materials were simpler than those of Roberts et al. (2007), involving direct (rather than indirect) object gaps. Thus, Hestvik et al.'s findings replicate for older children what Love and Swinney (2007) have already found for 4-to-6-year-olds using the alive/non-alive decision task, but they do not necessarily show that the picture-naming version of the task is to be preferred.

3. Behavioral methods for studying language production

While the adult psycholinguistic literature offers a range of different behavioral methods for investigating language comprehension, there are only a few experimental paradigms available that tap processes during language production, e.g., implicit priming (Roelofs 2002, among others), the picture-word interference paradigm (e.g., Schriefers, Meyer, & Levelt 1990), syntactic priming (e.g., Pickering & Branigan 1998), and speeded production (e.g., Prasada, Pinker, & Snyder 1990). Of these the latter two have been adapted to the study of children's language production. In the following, I will first briefly explain syntactic priming and then in some more detail the speeded production task.

3.1 Syntactic priming

When people produce sentences they are likely to maintain aspects of syntactic structure from one sentence to the next, a phenomenon that is called syntactic priming. The conditions under which syntactic priming occurs are thought to reveal aspects of grammatical encoding during production. In syntactic priming studies (see, e.g., Bock, Loebell, & Morey 1992; Pickering & Branigan 1998), subjects provide continuations for partial sentences of both prime fragments and target sentences. Prime fragments are such that the most likely completion is of a particular form; for example, for fragments such as (10a) and (10c) a completion with a prepositional object is highly likely. By contrast, the target fragments (10b) and (10d) end after the verb so that subjects have a choice between a prepositional and a double-object construction for their continuation.

(10) Priming of prepositional object construction
a. The rock star sold some cocaine … → … *to an undercover agent*
b. The girl handed … → … *the paintbrush to the man*

c. Mary baked a cake … → … *for her boss*
d. The girl handed … → … *the paintbrush to the man*

Priming of double-object construction
e. The rock star sold an undercover agent ... →... *some cocaine*
f. The girl handed … → … *the man the paintbrush*

Examining three-place predicates such as those in (10b), (10d), and (10f) as targets, Pickering and Branigan (1998) found priming effects caused by prepositional object constructions (10a) and (10c), and double-object constructions (10e) in adults. Thus, the completion of (10a) primed participants to complete (10b) using a prepositional rather than a double-object construction and vice versa for a prime such as (10e). Interestingly, it was found that while a prepositional phrase with *for* (10c) primed the prepositional object construction (with *to*) for *give* (10d), the same lexical item (*to*) in a different syntactic function (e.g., in *Mary brought a book to study*), does not prime the prepositional object construction for *give*. This finding suggests that production priming effects are abstract and syntactic in nature rather than purely based on lexical information. Syntactic priming in these cases has been explained in terms of phrase-structure rules. Thus, the construction of a prime sentence such as (10a) involves a rule that expands a VP into V+NP+PP. Once employed for the construction of the prime sentence, the rule may retain some residual activity when the target fragment is completed thus making an NP+PP completion more likely than an NP+NP completion.

Production priming has also been used with children, with modifications in the design. For example, Savage, Lieven, Theakston and Tomasello (2003) presented 4-to-6-year-old children with prime sentences spoken by the experimenter along with a prime picture. The child was then asked to repeat the prime sentence, e.g., an active or a passive sentence. Then, the child was presented with the target picture and asked 'What's happening here?' to examine whether the child was prompted to produce an active or a passive sentence depending on the prime sentence presented before. Savage et al. (2003, 2006) obtained priming effects for 6-year-olds, whereas for 4-year-olds priming effects only occurred in cases of high lexical overlap between primes and targets, i.e., when the prime sentence contained the same lexical verb as required for the target picture. Savage et al. took this to mean that 4-year-olds' representations of actives and passives are lexically-specific rather than abstract or syntactic in nature. On the other hand, Huttenlocher, Vasilyeva and Shimpi (2004) found a structural priming effect for 4-year-olds even when the prime and the target sentences did not share lexical content words. Here, I will not further discuss these conflicting results. It should be noted, however, that Savage et al.'s claims about the lack of abstract syntactic knowledge in children are only based on production measures, which may under-estimate a child's linguistic knowledge. Their view that 4-year-olds lack abstract syntactic knowledge would be more convincing if they had converging evidence from other sources, e.g., from comprehension measures.

For our present concerns, it is important to note that the syntactic priming technique as it stands is not time-sensitive as it does not provide any measure of the time-course of grammatical encoding. We may ask, for example, whether priming effects in this task unfold predictively and how this precisely happens over time; see Snedeker and Thothathiri (this volume) for ways of incorporating time-sensitive measures into the syntactic priming task.

3.2 Speeded production

The measurement of production latencies offers a way to examine automatic processes involved in children's spoken productions. In a speeded-production task, participants are asked to produce as quickly and accurately as possible a particular word form, e.g., an inflected form (*walked*) for an auditorily presented verb stem (*walk*). Accuracy rates and production latencies are measured, the latter of which provide the crucial on-line measure.

Several research teams have used this technique to examine potential processing differences between regular and irregular inflection in adults (Prasada et al. 1990; Ullman 1993; Beck 1997; Lalleman, van Santen, & van Heuven 1997; Buck-Gengler, Menn, & Healy 2004). The purpose of these studies was to determine to what extent the real-time production of an inflected word relies on lexical look-up, i.e., upon retrieval of whole word forms stored in memory, and to what extent it depends on computational processes of, for example, combining stems or roots with affixes (*walk+-ed*). The rationale is that if an inflected word form is stored as a whole, then retrieval should be faster for high-frequency than for low-frequency ones, and this contrast should be measurable in production latencies. This is a sensible assumption, since lexical retrieval and storage are known to be affected by a word's frequency. On the other hand, if regularly inflected forms are computed from their morphological constituents during production (rather than retrieved as whole word forms from memory), then the word frequency of a regularly inflected form (e.g., the frequency of *walked*) should not affect production latencies. Hence, of two regularly inflected forms that have the same stem frequency but differ with respect to their past-tense frequency (e.g., *jump* and *boil* which both have a stem frequency of 26 per million and past-tense frequencies of 32 for *jumped* and 1 for *boiled, see* Prasada et al. 1990), producing the one with the lower past-tense frequency should not take longer than the production of the high-frequency word form.

This paradigm has produced reasonably clear and replicable effects for adults. All the studies mentioned above found a frequency advantage for irregulars, i.e., shorter response times for high-frequency than for low-frequency irregulars, and no corresponding advantage for high-frequency forms amongst regulars (see Pinker 1999: 129ff. for review). Let us consider more closely a study (Clahsen, Hadler, & Weyerts 2004) in which the speeded production task was used to examine morphological processing in children.

3.2.1 *Investigating the production of inflected words*

Clahsen et al. (2004) examined regular and irregular participle forms of German with high and low frequencies in two age groups of children (5- to 7-year-olds, and 11- to 12-year-olds) and in a group of adult native speakers. Participle formation in German involves two suffixes *-t* and *-n*. All regular verbs are suffixed with *-t* (parallel to *-ed* in English), e.g., *kaufen-gekauft* 'buy-bought'; all irregular verbs have the ending *(-e)n* (*laden-geladen* 'load-loaded'), akin to what we still see in English for a small number of verbs such as *write-written* or *take-taken*. In addition, many German participles of both regular and irregular verbs have a prefix (*ge-*). Prefixation, however, is prosodically determined: *ge-* only occurs when the verbal stem is stressed on the first syllable. Since German verbal stems are often stressed on the first syllable, the (unstressed) *ge-* prefix is highly frequent.

In our experiment, participants listened to stem forms of verbs presented in a sentential context, and were asked to produce corresponding participle forms as quickly and accurately as possible. To make the experiment more appealing to children, pictures were presented along with spoken sentence fragments such as those in (11). The first picture for each fragment depicted the subject of the sentence (e.g., *der Frosch* 'the frog') and appeared in the upper left-hand corner of the screen at the same time at which the subject NP was heard. Then, participants listened to an auxiliary (e.g., *hat* 'has') followed by the object NP (e.g., *die Fliege* 'the fly'), at which time the corresponding picture appeared in the bottom right-hand corner of the screen. Then, a new screen was presented with a cartoon figure that moved its lips while producing a verb stem with rising intonation so as to indicate that it was not sure what the correct word form might be. Participants were told that the cartoon figure does not know German very well and that the participants' task was to help out and to provide the correct form as quickly and as accurately as possible.

(11) a. Der Frosch hat die Fliege … fress?
'The frog has the fly … eat?'

b. Der Mann hat die Treppe … wisch?
'The man has the stairs … mop?'

Table 3. Mean production latencies (in ms.) for high and low-frequency (regular and irregular) participle forms (adapted from Clahsen et al. 2004).

	Adults (N = 35)	**5- to 7-Year-Olds** (N = 20)	**11- to 12-Year-Olds** (N = 20)
Irregular-High	947	1223	1078
Irregular-Low	1002	1283	1130
Regular-High	958	1257	1088
Regular-Low	947	1188	1049

The results from this study, presented in Table 3, can be summarized in three points. Firstly, while adults produced hardly any morphological errors in participle formation, children were found to overregularize the regular *-t* suffix to verbs that require the irregular *-n* suffix, with higher error rates on low than on high-frequency irregulars. By contrast, over-applications of the irregular suffix to regular verbs were extremely rare (less than 1%). Secondly, the overall production latencies were found to decrease with age, 5- to 7-year-olds having a mean production latency of 1238 ms, 11- to 12-year-olds of 1086 ms, and adults of 963 ms. Thirdly, whereas all participant groups had shorter production latencies for high-frequency irregulars than for low-frequency ones, both age groups of children showed a reverse frequency effect for regulars, longer production latencies for high than for low-frequency regulars. This contrast was more pronounced for the 5- to 7-year-old children than for the 11- to 12-year-olds.

These results can be taken to indicate that two mechanisms for morphological processing, lexical storage and morphological computation, are employed by children as well as by adults but that lexical access is less efficient for children. Overregularization errors arise when access to the lexical entry of an irregular form fails. Consequently, children produce more of such errors than adults. Children took longer to produce participles than adults, another indication of less efficient lexical access. Finally, reverse-frequency effects arise from the retrieval of stored high-frequency regulars that inhibit morphological computation (Pinker 1999). Hence the production of high-frequency regulars involves memory access, and this interferes with morphological computation (which is available for both high and low-frequency regulars) in that it slows down the production of high-frequency regulars relative to low-frequency ones for which morphological computation is not impeded by any stored forms. Slow lexical retrieval increases this contrast, hence the decrease of the reverse frequency effect from the younger to the older child groups.

3.2.2 *Methodological issues*

The design of materials for an experiment of this kind requires careful consideration of a number of potentially confounding factors. One concern is that phonetic differences in onset length may affect production latencies. For example, stops are intrinsically shorter in duration than fricatives. Intrinsic segmental duration differences will affect production latencies, because these are measured from the onset of the stimulus. Compare, for example, the release of the initial consonant in *t*ea versus the beginning of turbulence in the initial consonant in *s*ea. Moreover, measuring the precise onset of words beginning with, for example, fricatives or nasals is more difficult than measuring those beginning with stops. These potential confounds can be addressed in the materials design, which is precisely what we did in Clahsen et al. (2004). All experimental items were participle forms that required *ge*-prefixation, thus precluding any effects of onset length on production latencies. At the same time, none of the filler items required *ge*- (e.g., *verlieren* – *verloren* 'lose – lost') thus making sure that a participle form could only be produced after the presentation of a particular verb stem. It is also necessary to control the duration of the verb stems presented as stimuli for elicitation and to ensure that they do not differ across experimental conditions. This is because production latencies are determined by measuring the time lag between the onset of the stem form given to participants and the onset of their response, and obviously, differences in stem durations would obscure these measurements.

Another concern is that the critical items are elicited as the final words of previously presented sentential fragments, as illustrated in (11), and that depending on the contents of the sentence fragments, it might be possible to anticipate or guess the final word of the sentence before encountering the verb stem provided. To address this possibility, the materials need to be pre-tested to make sure that the critical items are equally unpredictable from the sentence fragments chosen for the speeded production task across the various experimental conditions.

Furthermore, care needs to be taken to ensure that the sentence fragments plus verb stems presented to participants do not sound unnatural. The materials for the Clahsen et al. study, for example, were read as complete sentences together with the corresponding participle and pre-recorded digitally. The verb stems were separately read and also recorded. Sentence fragments were cut off before the onset of the participle. To make sure that the stimuli sounded as natural as possible, the audio files containing the verb stems were inserted at exactly the same point at which the participles were cut out from the complete sentences. Thus, the time lag between a sentence fragment and a verb stem was identical to the lag between the participle and the preceding word in the complete sentences.

Finally, and most importantly, the critical items need to be selected according to their frequency in relevant corpora. For studies with young (pre-literate) children, frequency information should be gathered from corpora of spoken language and ideally from corpora of child-directed speech, because frequency dictionaries or corpora of written language (e.g., newspapers) may contain words unfamiliar to young children.

Summarizing, the speeded production task provides an efficient measure of processes involved in language production. The technique offers a time-sensitive measure, even though the response latencies do not only reflect production processes but also include the time needed for recognizing the verb stem presented. Stimulus presentation is fairly natural requiring normal listening, but various potentially confounding factors (mentioned above) need to considered to avoid artifacts. The task is not demanding and, in the modified version (Clahsen et al. 2004), appropriate for children above the age of 5. Indeed, none of the children tested found the task particularly difficult, and most of them enjoyed the experiment. Moreover, the task requires minimal technical apparatus (a PC and a microphone) and can be performed in any quiet room, even outside the research laboratory. As regards its linguistic versatility, the speeded-production task seems to be well-suited to examine word-level processing, e.g., the production of inflected word forms of bare stems, plurals from singulars, etc., but less so for studying sentence-level processing.

4. Summary

The three behavioral methods we focused on, self-paced reading and listening, cross-modal priming, and speeded production, all provide time-sensitive measures, an essential requirement for studying on-line processing. Unfortunately, however, each of these measures has its limitations. Self-paced reading and listening have a relatively low temporal resolution (compared, for example, to ERPs and eye movement measures) due to the way the stimuli are presented. In cross-modal priming experiments, response times are only measured at specific test points in a sentence thus providing a snapshot view of the state of the language processor at these points rather than a continuous measure of on-line sentence processing. The response latencies that are measured in the speeded-production task include the recognition times required for the auditory stimuli and cannot be taken as a pure measure of language production processes.

Clearly, eye tracking and ERPs provide better measures of the time-course of processing, but the behavioral methods discussed in this chapter will no doubt have a place in future research in this field, and this is for a number of reasons. Firstly, any psycholinguistic technique (including eye tracking and ERPs) has its limitations and is in danger of producing artifacts, e.g., due to an experiment's specific task demands. One way around this problem is to find converging evidence from other sources, e.g., by replicating an effect seen with one technique with a different technique. Behavioral techniques can be useful for this purpose. Secondly, behavioral techniques require relatively little technical equipment and can be administered without bringing children into the research laboratory. This makes them ideally suited for piloting experimental designs, and for working with populations in out-of-reach places. Thirdly, compared to, for example, ERP experiments which require many items per condition due to signal averaging, the behavioral methods mentioned above require fewer critical items and can typically be administered within a single experimental session, thus avoiding potential artifacts such as those caused by training effects and fatigue. Finally, unlike ERP or eye tracking experiments for which an electrode cap or a head-band needs to be attached to the child, behavioral techniques do not require any direct physical contact with a participant. It will therefore be much easier to get ethical approval for behavioral experiments than for any technique involving physiological measures, and there may be circumstances in which such considerations are a decisive factor.

Acknowledgements

I am grateful to Sonja Eisenbeiss, Claudia Felser, Eva Fernández, Ted Gibson, Theodore Marinis, Leah Roberts, and Matt Traxler for comments on an earlier version of this chapter.

References

Beck, M.-L. (1997). Regular verbs, past tense and frequency: Tracking down a potential source of NS/NNS competence differences. *Second Language Research*, *13*, 93–115.

Bever, T. G. & McElree, B. (1988). Empty categories access their antecedents during comprehension. *Linguistic Inquiry, 19*, 35–43.

Bever, T. & Sanz, M. (1997). Empty categories access their antecedents during comprehension. Unaccusatives in Spanish. *Linguistic Inquiry, 28*, 68–91.

Bock, K., Loebell, H. & Morey, R. (1992). From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological Review, 99*, 150–171.

Booth, J., Perfetti, C., & MacWhinney, B. (1999). Quick, automatic and general activation of orthographic and phonological representation in young readers. *Developmental Psychology*, *35*, 3–19.

Booth, J., MacWhinney, B., & Harasaki, Y. (2000). Developmental differences in visual and auditory processing of complex sentences. *Child Development*, *71*, 981–1003.

Buck-Gengler, C., Menn, L., & Healy, A. (2004). What 'mice trap' tells us about the mental lexicon. *Brain and Language, 90*, 453–464.

Carreiras, M. & Clifton, C. (1993). Relative clause interpretation preferences in Spanish and English. *Language and Speech*, *36*, 353–372.

Carreiras, M. & Clifton, C. (1999). Another word on parsing relative clauses: Eyetracking evidence from Spanish and English. *Memory and Cognition*, *27*, 826–833.

Clahsen, H. & Featherston, S. (1999). Antecedent priming at trace positions: Evidence from German scrambling. *Journal of Psycholinguistic Research*, *28*, 415–437.

Clahsen, H., Hadler, M., & Weyerts, H. (2004). Speeded production of inflected words in children and adults. *Journal of Child Language*, *31*, 683–712.

Carreiras, M. & Clifton, C. (1999). Another word on parsing relative clauses: Eyetracking evidence from Spanish and English. *Memory and Cognition*, *27,* 826–833.

Clifton, C. & Frazier, L. (1989). Comprehending sentences with long-distance dependencies. In G. M. Carlson & M. K. Tanenhaus (Eds.), *Linguistic structure in language processing*. Dordrecht: Kluwer.

Cuetos, F. Mitchell, D., & Corley, M. (1996). Parsing in different languages. In M. Carreiras, J. García-Albea, & N. Sebastián-Gallés (Eds.), *Language processing in Spanish* (pp. 145–187). Mahwah NJ: Lawrence Erlbaum Associates.

De Vincenzi, M. & Job, R. (1993). Some observations on the universality of the late-closure strategy. *Journal of Psycholinguistic Research*, *22*, 189–206

Felser, C., Marinis, T., & Clahsen, H. (2003). Children's processing of ambiguous sentences: A study of relative clause attachment. *Language Acquisition*, *11,* 127–163.

Fernández, E. M. (2003). *Bilingual sentence processing,* Amsterdam: John Benjamins.

Ferreira, F., Henderson, J. Anes, M. Weeks, P, & McFarlane, D. (1996). Effects of lexical frequency and syntactic complexity in spoken language comprehension: Evidence from the auditory moving window technique. *Journal of Experimental Psychology: Learning, Memory and Cognition, 22,* 324–335.

Ferreira, F., Anes, M., & Horine, M. (1996). Exploring the use of prosody during language comprehension using the auditory moving window technique. *Journal of Psycholinguistic Research*, *25,* 273–290.

Frazier, L. & Clifton, C. (1996). *Construal.* Cambridge MA: The MIT Press.

Gathercole, S., & Baddeley, A. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language, 28,* 200–213.

Gibson, E., Pearlmutter, N., Canseco-Gonzalez, E., & Hickock, G. (1996). Recency preferences in the human sentence processing mechanism. *Cognition*, *59,* 23–59.

Gibson, E. & Warren, T. (2004). Reading-time evidence for intermediate linguistic structure in long-distance dependencies. *Syntax*, *7,* 55–78.

Gilboy, E., Sopena, J., Clifton, C., & Frazier, L. (1995). Argument structure and association preferences in Spanish and English compound NPs. *Cognition, 54,* 131–167.

Gordon, P. C., Hendrick, R., & Foster, K. L. (2000). Language comprehension and probe–list memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *26,* 766–775.

Hestvik, A., Schwartz, R., Tornyova L., & Datta, H. (2005). Picture-naming shows children reactivate antecedents at trace positions. Poster presented at the Generative Approaches to Language Acquisition (GALA) Conference, Siena, Italy.

Huttenlocher, J., Vasilyeva, M., & Shimpi, P. (2004). Syntactic priming in young children. *Journal of Memory and Language, 50*, 182–195.

Just, M. A., Carpenter, P. A., & Wooley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General, 111,* 228–238.

Kail, M. & Diakogiorgi, K. (1998). On-line integration of morpho-syntactic cues by Greek children and adults: A cross-linguistic perspective. In N. Dittmar & Z. Penner (Eds.), *Issues in the theory of language developmen*t. Bern: Peter Lang.

Kail, M. (2004). On-line grammaticality judgments in French children and adults: A crosslinguistic perspective. *Journal of Child Language, 31*, 713–737.

Kidd, E. & Bavin, E. (2007). Lexical and referential influences on on-line spoken language comprehension: A comparison of adults and primary school age children. *First Language, 27*(1), 29–52.

King, J. & Kutas, M. (1995). Who did what and when? Using word- and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience, 7,* 376–395.

Lalleman, J., van Santen, A., & van Heuven, V. (1997). L2 processing of Dutch regular and irregular verbs. *Review of Applied Linguistics*, *115/116*, 1–26.

Love, T. & Swinney, D. (1996). Coreference processing and levels of analysis in object-relative constructions: Demonstration of antecedent reactivation with the cross-modal priming paradigm. *Journal of Psycholinguistic Research, 20*, 5–24.

Love, T. & Swinney, D. (2007). The processing of non-canonically ordered constituents in long distance dependencies by pre-school children: A real-time investigation. *Journal of Psycholinguistic Research, 36*, 191–206.

MacDonald, M. (1989). Priming effects from gaps to antecedents. *Language and Cognitive Processes, 4*, 35–56.

Mazuka, R. (1998) *The development of language processing strategies. A cross-linguistic study between Japanese and English*. Hillsdale NJ: Lawrence Erlbaum Associates.

McElree, B. & Bever, T. (1989). The psychological reality of linguistically defined gaps. *Journal of Psycholinguistic Research, 18,* 21–35.

McElree, B. & Griffith, T. (1995). Syntactic and thematic processing in sentence comprehension: Evidence for a temporal dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 134–157.

McElree, B. & Griffith, T. (1998). Structural and lexical constraints on filling gaps during sentence comprehension: A time-course analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 432–460.

McKee, C. (1996). On-line methods. In D. McDaniel, C. McKee & H. Smith Cairns (Eds.), *Methods for assessing children's syntax* (pp. 189–208). Cambridge MA: The MIT Press.

McKee, C., Nicol, J., & McDaniel, D. (1993). Children's application of binding during sentence processing. *Language and Cognitive Processes, 8,* 265–290.

McKoon, G. & Ratcliff, R. (1994). Sentential context and on-line lexical decision. *Journal of Experimental Psychology: Learning, Memory and Cognition, 20*, 1239–1243.

McKoon, G., Ratcliff, R. & Allbritton, D. (1996). Sentential context effects on lexical decisions with a cross-modal instead of an all-visual procedure. *Journal of Experimental Psychology: Leaning, Memory, and Cognition, 22*, 1494–1497.

Mitchell, D. C. (2004). On-line methods in language processing: Introduction and historical review. In M. Carreiras & C. Clifton (Eds.), *The on-line study of sentence comprehension: Eyetracking, ERP and beyond* (pp. 15–32). Hove: Psychology Press.

Miyamoto, E. & Takahashi, S. (2001). Antecedent reactivation in the processing of scrambling in Japanese. *MIT Working Papers in Linguistics, 43,* 127–142.

Nakano, Y., Felser, C., & Clahsen, H. (2002). Antecedent priming at trace positions in Japanese long-distance scrambling. *Journal of Psycholinguistic Research, 31*, 531–571.

Nicol, J. & Swinney, D. (1989). The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research, 18*, 5–20.

Nicol, J., Swinney, D., Love, T., & Hald, L. (2006). The on-line study of sentence comprehension: Examination of dual-task paradigms. *Journal of Psycholinguistic Research*, *35*, 215–231.

Pickering, M. & Branigan, H. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language*, *39,* 633–651.

Pickering, M. (1993). Direct association and sentence processing: A reply to Gibson & Hickok. *Language and Cognitive Processes*, *8,* 163–196.

Pinker, S. (1999). *Words and rules. The ingredients of language*. New York NY: Basic Books.

Prasada, S., Pinker, S., & Snyder, W. (1990). Some evidence that irregular forms are retrieved from memory but regular forms are rule-generated. Paper presented at the 31st Annual Meeting of the Psychonomic Society, New Orleans, LA.

Roberts, L., Marinis, T., Felser, C., & Clahsen, H. (2007). Antecedent priming at trace positions in children's sentence processing. *Journal of Psycholinguistic Research*, *36*(2), 175–188.

Roelofs, A. (2002). Storage and computation in spoken word production. In S. Nooteboom, F. Weerman, & F. Wijnen (Eds.), *Storage and computation in the language faculty* (pp. 183–216). Dordrecht: Kluwer.

Savage, C., Lieven, E., Theakston, A., & Tomasello, M. (2003). Testing the abstractness of children's linguistic representations: Lexical and structural priming of syntactic constructions in young children. *Developmental Science, 6*, 557–567.

Savage, C., Lieven, E., Theakston, A., & Tomasello, M. (2006). Structural priming as implicit learning in language acquisition: The persistence of lexical and structural priming in 4-year-olds. *Language Learning and Development, 2*, 27–49.

Schriefers, H., Meyer, A., & Levelt, W. (1990). Exploring the time course of lexical access in language production: Picture-word interference. *Journal of Memory and Language, 29*, 86–102.

Schütze, C. (1996). *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. Chicago IL: University of Chicago Press.

Swets, B, Desmet, T., Hambrick, D. Z., & Ferreira, F. (2007). The role of working memory in syntactic ambiguity resolution: A psychometric approach. *Journal of Experimental Psychology: General*, *136*(1), 64–81.

Traxler, M. (2002). Plausibility and subcategorization preference in children's processing of temporarily ambiguous sentences: Evidence from self-paced reading. *Quarterly Journal of Experimental Psychology*, *55A,* 75–96.

Traxler, M. (2005). Plausibility and verb subcategorization in temporarily ambiguous sentences: Evidence from self-paced reading. *Journal of Psycholinguistic Research*, *34,* 1–30.

Traxler, M. & Pickering, M. (1996). Plausibility and the processing of unbounded dependencies: An eye-tracking study. *Journal of Memory and Language, 35*, 454–475.

Traxler, M, Pickering, M., & Clifton, C. (1998). Adjunct attachment is not a form of lexical ambiguity resolution. *Journal of Memory and Language, 39*, 558–592.

Tyler, L. & Marslen-Wilson, W. (1981). Children's processing of spoken language. *Journal of Verbal Learning and Verbal Behavior, 20*, 400–416.
Ullman, M. T. (1993). The computation of inflectional morphology. PhD dissertation, MIT.
Wulfeck, B. (1993). A reaction time study of grammaticality judgments in children. *Journal of Speech and Hearing Research, 36,* 1208–1215.

CHAPTER 2

Event-related brain potentials as a window to children's language processing

From syllables to sentences

Claudia Männel and Angela D. Friederici

The present paper gives an overview of our recent research on the neurocognition of language acquisition. Our research aims to gain a more detailed understanding of the developmental stages of the language acquisition process and its underlying brain mechanisms. Here, we utilize the method of event-related brain potentials, which has revealed specific electrophysiological indices for various aspects of language processing in adults. These electrophysiological parameters can serve as templates to define the hallmarks of language acquisition. The research presented demonstrates that the method of event-related brain potentials is a powerful tool to investigate and monitor early stages of language acquisition and provides further insights into the neural correlates of language processing in infants and children.

1. Introduction

The wonder of language acquisition, with its remarkable speed and high levels of success, remains a mystery. At birth, infants are able to communicate their basic needs by different ways of crying. Also, from birth on, infants show a preference for the sound of their native language. Following these first language-related steps there is a fast progression in the development of perceptive and expressive language skills. At around 4 months, babies start to babble, the earliest stages of language production. A mere twelve months after birth most babies start to speak their first words, and about half a year later even short sentences. Finally, at the end of most children's third year of life, they have acquired at least 500 words and know how to combine them into meaningful utterances. Thus, they have mastered the entry into their native language: they have acquired a complex system with the typical sounds of a language, these sounds are combined in different ways to make up a large vocabulary, and the vocabulary entries are related together by means of syntactic rules. All of this occurs at an amazing speed and to some extent independently of the environmental conditions.

Although a quite detailed outline of the language acquisition process exists (see Bee & Boyd 2007; Clark 2003; Klann-Delius 1999; Szagun 2006), many questions remain. For example, how do children actually learn their mother tongue so easily and what are the exact developmental stages? Does language acquisition happen in a continuous manner or a discontinuous manner? Here, the *continuity hypothesis* of language acquisition holds that language processes change quantitatively over the course of the child's development but that the processes are in principle present from early on and in a similar form to the way they are present in adults. In contrast, the *discontinuity hypothesis* proposes qualitative rather than quantitative changes during the child's development until an adult-like status is reached. In this context, then, what types of developmental changes occur in the brain during language development? In other words, what is the neurophysiological basis of the various processes and steps in language acquisition? Studying how children so readily acquire language is not easily accomplished, because a good deal of learning takes place before the child is able to speak and show overt responses to what he or she actually perceives. For example, it is not an insignificant methodological challenge to develop ways of asking children whether they know a word before they can demonstrate a verbal reaction to this question. Children's perceptive language skills develop much earlier than their expressive skills but are for the most part beyond the scope of observation.

The method of event-related brain potentials (ERP) allows us to virtually look into the brain, where the acquisition of language is taking place, during the course of acquisition. The use of the ERP method to investigate on-line cognitive processes in adults has been successfully proven. In adults, there are particular ERP components that appear to be specific to various aspects of language processing. We also use this method to study language acquisition in infants over the course of the first three years of life since this method inherits several advantages that become apparent specifically in the work with infants.

In this chapter, we demonstrate that the ERP method is a powerful tool to investigate and monitor early stages of language acquisition. Specifically, we show that the ERP method delivers information about the neural correlates of language processes and therefore provides a better understanding of the *how* and *when* of the developmental stages in the language acquisition process. First, we explain the method at hand thoroughly (some details about its successful application in work with infants are provided in the Appendix). Then, we briefly outline the research field as well as sketch the developmental stages in language acquisition and their associated ERP components. Then, we give an overview of our own ERP research on the different aspects of language acquisition during infants' first three years of life. Thus, we describe, by means of ERP, the processing of phonological, semantic, and syntactic information in infants and children and discuss these results in the light of other neurophysiological and behavioral studies.

2. The method of event-related brain potentials

2.1 What is measured? EEG and ERP

Electroencephalography (EEG) is a non-invasive method used to measure voltage fluctuations on the scalp's surface. These voltage fluctuations comprise summed post-synaptic electric potentials generated by similarly aligned and simultaneously firing pyramidal cells in the neocortex (for more detail see Lopes da Silva 1991; Speckman & Elger 1993). Consequently, the recorded electric activity reflects a wide range of brain functions, including various states of activation, relaxation, tiredness, engagement in cognitive tasks, etc. EEG data recorded for a certain time period contain background activity as well as changes of electric activity in response to single events, such as words and sentences. In response to those events, the recorded brain signal can be broken up into its frequency bands, i.e., regular patterns of electric potentials in a defined time window (Fourier analysis). The recorded signal can also be analyzed for voltage fluctuations that are time-locked to sensory, motor, and specifically cognitive events, so-called event-related brain potentials (ERPs; for reviews see Coles & Rugg 1995; Fabiani, Gratton, & Coles 2000; Regan 1989). These event-related voltage changes are relatively small compared to the ongoing background EEG activity. For this reason, the interpretation of EEG raw data in relation to single events is, if not impossible, at least very difficult (although there has been some advancement in single-trial analysis; see for instance Bansal, Sun, & Sclabassi 2004; Holm, Ranta-Aho, Sallinen, Karjalainen, & Müller 2006; Jung et al. 2001). To overcome this problem, repeated presentation of the stimulus of interest (e.g., a spoken sentence) and subsequent averaging are required. In order to analyze the recorded signal on a stimulus- or event-related basis, time-locked epochs of interest (for instance triggered by the onset of a spoken sentence) are extracted from the EEG raw data. To obtain a high signal-to-noise ratio a minimum number of trials (e.g., 30 sentences of the same sentence type) is required for averaging. Prior to averaging, optional filtering and eye movement correction additionally help to remove artifacts and other unrelated brain activity. Data processing and subsequent trial averaging produce a smooth curve of changes in electric activity over time since they eliminate or at least reduce the effect of random noise distributed by each of the single trials.

The resulting wave form represents the average processing of a stimulus over time and consists of a sequence of positive and negative voltage fluctuations, referred to as components, waves, deflections, or peaks. These event-related components are associated with various sensory, motor, and cognitive processes and reflect covert and overt information processing. The N1 component, a negative fluctuation around 100 ms, for example, reflects early sensory responses and was found to be modulated by attention (e.g., Hillyard, Vogel, & Luck 1998; Mangun 1995; Woldorff et al. 1993).

In sum, EEG measurement and subsequent averaging of stimulus-triggered epochs to gain ERPs provide a direct, non-invasive measure of the temporal course of changes in electric activity that directly relate to neuronal information processing.

2.2 Why should we use ERPs? Advantages and disadvantages of the ERP method

Behavioral experiments measure overt responses such as reaction time and number of correct answers. These measures provide us with the end product of the cognitive processes engaged during the perception and evaluation of a given stimulus. However, those techniques are not capable of monitoring the actual on-line cognitive processes that lead to the observed behavior. Although eyetracking methods may deliver on-line parameters of the ongoing information processing, these measures are nevertheless indirect indicators of the underlying neuronal mechanisms. In contrast, methods in the field of Cognitive Neuroscience – such as neuroimaging techniques and the ERP method – inform us about the on-line stages of information processing in the brain.

The ERP method features excellent temporal resolution, as it delivers information in millisecond accuracy about the time course of brain responses. In this way ERPs provide a mental chronometry, i.e., an exact temporal sequencing of information processing. In comparison to neuroimaging techniques (e.g., *fMRI* and *PET*), the spatial resolution for the identification of the neural generators of the obtained signal is relatively poor in ERPs, since maximal amplitude measures at certain electrode sites only provide information about where neural activity, evoked by certain stimuli, arrives at the scalp's surface. However, there are some sophisticated source localization techniques that estimate the neural generators based on the measured electric scalp potential. There are two categories of source localization methods: techniques that postulate distributed current sources (e.g., the minimum norm-based technique LORETA; Pascual-Marqui 2002; Pascual-Marqui, Michel, & Lehmann 1994) and techniques that assume equivalent current dipoles as neural origins (e.g., the BESA technique; Scherg, Vajsar, & Picton, 1989; Scherg & von Cramon 1986). One of the few developmental studies that applied source localization investigated age-related changes in the auditory system based on dipole source modeling (Ponton, Eggermont, Khosla, Kwong, & Don 2002). Independent of the applied localization technique, additional data from functional imaging and clinical studies considerably help to constrain source analysis. In the field of developmental cognitive neuroscience, however, this information is mainly restricted to adult studies (for an overview on fMRI research on language processing see for instance Friederici 2004a), since there are only few systematic children fMRI studies (e.g., Brauer & Friederici 2007) and investigations on aphasia in children (for an overview see Friederici 1994).

In working with infants and children, when investigating the question of how children actually acquire their mother tongue and how their language processing abilities develop over time, the advantages of the ERP method become readily apparent. For EEG recordings, no overt responses are necessary since EEG directly measures brain activity to stimuli, thus considerably facilitating work with infants. The fact that ERP components are direct indicators of the underlying brain processes implies not only that there are no task requirements for measurement but also that brain processes evoked by certain stimuli might be detectable before there is a behavioral correspondence observable at a certain stage in the child's development. Although behavioral methods used in infant research, e.g., the *headturn paradigm*, the *preferential looking paradigm* and *eyetracking techniques* (see chapters in this volume by Clahsen; Fernald, Zangl, Portillo, & Marchman; Trueswell; Snedeker & Thothathiri), require a less complicated set-up and can be performed in a more natural setting, these methods are at the same time more prone to external interferences. Regarding neuroimaging techniques, there are still some limitations in the work with infants and young children (but see Redcay, Haist, & Courchesne 2006; see also Hebden 2003; Meek 2002 on optical imaging in infants). First, movement restrictions during brain scanning make it rather difficult to work with children. Second, there is still an ongoing discussion whether the BOLD signal in adults is comparable to the one in children and whether the applied adult models are appropriate for infant research (for discussion see Anderson et al. 2001; Marcar, Strassle, Loenneker, Schwarz, & Martin 2004; Martin et al. 1999; Rivkin et al. 2004; Schapiro et al. 2004).

In addition to considering how a method differs from others and how practical it might be, it is also important to consider how different methods deliver different kinds of information. The decision to use a specific method highly depends on the kind of information sought: the neuronal correlates of information processing in their spatial and/or temporal resolution or the behavioral consequences that follow from these processes. Ideally, a combination of various measures using the complementary abilities of different methods should be sought.

2.3 What do we get in the end? ERP components and their interpretation

The components of event-related brain potentials can be described in terms of their amplitude/polarity, their latency, and their characteristic scalp distribution/topography. *Amplitude* indicates the extent to which a response to an experimental stimulus is elicited, i.e., the amount of neural activity. Dependent on the pole orientation of the measured electric field, neural activity is reflected in positive and negative deflections, indicating neuronal discharging and charging, respectively. *Latency* refers to the point in time at which this activation occurs relative to stimulus onset. Given these two parameters, specific components are labeled as follows: waves with a negative-going deflection are designated by *N*, waves with a positive-going deflection by *P*, while the time (in milliseconds) from stimulus onset to certain wave peaks is indicated by a number. The N100 component, for example, refers to a negativity occurring at about 100 ms after stimulus onset. Note, however that components are often labeled according to the order of their occurrence during stimulus processing (e.g., N1, P2, N2) and their functional significance rather than their polarity and latency parameters per se. For instance, the P300 component is known to occur in various oddball paradigms in response to deviant stimuli and reflects memory- and context-updating processes after stimulus evaluation (Donchin & Coles 1988). Dependent on stimulus discrimination difficulty, stimulus complexity, and categorization demands the P300 latency varies between 300 ms and 700 ms post-stimulus (e.g., Katayama & Polich 1998; Daffner et al. 2000). *Scalp distribution* or *topography* provides information about a component's voltage gradient over the scalp at any point in time and therefore some information about the underlying neuroanatomical activity. As pointed out, conclusions from ERP data about the sources of neural generators can only be drawn in a restricted way when relying on the topographic information alone, e.g., lateralization to one hemisphere or distribution over posterior brain regions. The label of a component can also include information about its topography, e.g., ELAN for Early Left Anterior Negativity. Furthermore, some labels of ERP components depict the particular experimental paradigm in which they are evoked, e.g., MMN for Mismatch Negativity.

ERP components are considered to be indicators of the progression of information processing over time. Earlier components (up to about 100–200 ms after the onset of the stimulus) reflect essentially automatic processes and are modulated by the physical properties of a stimulus, such as the loudness and pitch of a spoken word. Later components are thought to reflect higher-order cognitive processing that is influenced by a person's intentions and actions, such as those present during a discrimination task between words and non-words. As pointed out, ERP components defined by specific parameters are likely to reflect different brain mechanisms engaged in stimulus processing. Changes in the parameters of a specific component indicate changes in the underlying cognitive mechanisms. For instance, a prolonged latency might point to a slowing down of a specific cognitive process and a reduced amplitude to a reduction in the processing demands or efficiency. Thus, ERP components are usually interpreted with respect to both their underlying neural mechanisms and their functional significance.

The interpretation of ERP components in infants demands some additional consideration. When dealing with infant ERPs, researchers should keep in mind the enormous physiological changes of the developing brain, concerning synaptic density, myelination, skull thickness, and fontanel state (see for example Mrzljak, Uylings, Van Eden, & Judas 1990; Pujol et al. 2006; Uylings 2006). For instance, the reduced synaptic density results in greater slow wave activity, possibly explaining why infant ERPs do not exhibit as many well-defined peaks as adult ERPs (Nelson & Luciana 1998). Infant ERPs usually show larger amplitudes than adult data, possibly due to difference in skull thickness, and longer latencies than adult ERPs, which, however, gradually decrease with increasing age (e.g., Jing & Benasich 2006; Kushnerenko et al. 2002). Thus, when comparing ERP components across age groups one should consider these maturational changes. Here, paradigms used in infant ERP experiments should be conducted in adults as well, thus achieving a target adult ERP pattern against which developmental comparisons can be made.

2.4 What does this tell us about language? ERP components related to language processing

In the domain of language processing, there are at least five functionally different components that reflect phonological, semantic, and syntactic processing in adults. The following components have been observed in ERP studies on language processing: (1) the Mismatch negativity (MMN), a negativity that occurs at around 100–250 ms post-stimulus and indicates the discrimination of acoustically or phonetically different stimuli (e.g., Näätänen 1990; Opitz, Mecklinger, Cramon, & Kruggel 1999) and is also modulated by language experience (Winkler et al. 1999); (2) the N400, a centro-parietally distributed negativity at around 400 ms post-stimulus that reflects lexical-semantic processes at both word level (e.g., Holcomb & Neville 1990) and sentence level (e.g., Kutas & Hillyard 1980, 1983); (3) the E/LAN, a left anterior negativity at around 150–350 ms post-stimulus, which occurs for on-line syntactic and morphosyntactic processes (e.g., Friederici, Pfeifer, & Hahne 1993; Hahne & Friederici 1999); (4) the P600, a centro-parietally distributed positivity at around 600 ms post-stimulus, which is related to processes of syntactic reanalysis and repair (e.g., Friederici & Mecklinger 1996; Kaan, Harris, Gibson, & Holcomb 2000; Osterhout & Holcomb 1992); and (5) the Closure Positive Shift (CPS), a centrally distributed positivity that has been observed in response to the processing of intonational phrases (e.g., Pannekamp, Toepel, Alter, Hahne, & Friederici 2005; Steinhauer, Alter, & Friederici 1999).

Over the last decades, a vast amount of studies in adults has demonstrated that the processing of different aspects of linguistic information can be clearly distinguished by means of these different ERP components. For a detailed description of the single components and the particular experimental conditions in which they are evoked we refer the reader to recent reviews (Friederici 2002, 2004b; Kutas & Federmeier 2000).

In infants and young children, recent neurophysiological research has demonstrated that most ERP components associated with phonological, semantic, and syntactic processes are quite similar to the ones observed in adults. This fact indicates that there are quantitative rather than qualitative changes in the language processes, reflected by particular ERP components, during infants' and children's development until an adult-like status is achieved. As aforementioned, ERP components in infants and children often show longer latencies and larger amplitudes as compared to the ones in adults, with a gradual latency and peak decrease as age increases (for comparisons between children and adult ERP data in language processing see for instance Hahne, Eckstein, & Friederici 2004; Oberecker, Friedrich, & Friederici 2005).

3. Developmental stages in language acquisition and their associated ERP components

In the following, we give an overview of ERP research on the different landmarks of language acquisition during a child's first three years of life, exemplified with experiments from our laboratory. We describe the processing of phonological, semantic, and syntactic information in infants and children and discuss these results in the context of other neurophysiological and behavioral studies. Figure 1 shows the developmental stages of auditory language perception and their associated ERP components. The developmental stages can be considered as interrelated phases during which new information is derived and processed based on previously acquired knowledge. From early on infants are able to discriminate speech

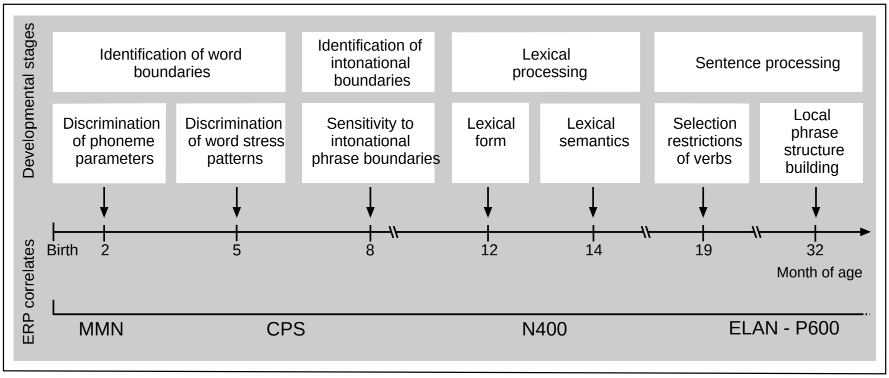


Figure 1. Developmental stages of language acquisition and their associated ERP components (modified from Friederici 2005).

sounds and word stress patterns, which facilitates the identification of content and function words in the sentential context. Furthermore, infants' early ability to process prosodic information at the sentence level, present in intonational phrase boundaries, aids the detection of syntactic boundaries. These processes eventually allow children to derive syntactic rules from speech input and provide the basis for the building of local syntactic structures and interphrasal relationships. The time course of the outlined developmental stages is based on the available ERP literature in infant research and is therefore only an approximation of the actual time course in language development. Thus, our overview should be understood as an attempt to sketch the language acquisition process based on the current knowledge within the framework of developmental cognitive neuroscience.

3.1 Processing of phonological/prosodic information

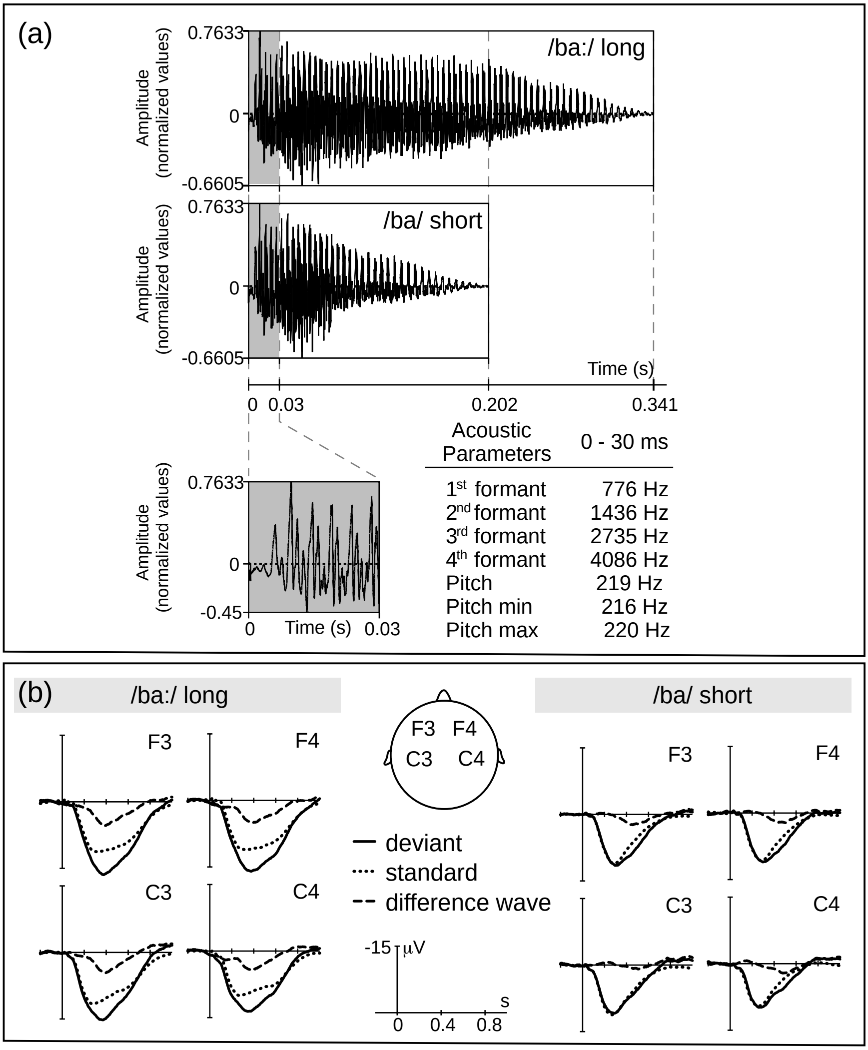
In general, infants' first steps into language are based on prosodic information. From birth on, prosodic cues facilitate the segmentation of the incoming speech stream into structural elements and therefore support the acquisition of lexical and syntactic units and eventually the derivation of syntactic ordering principles, a process called *prosodic bootstrapping*. Infants' first challenge is to extract the phonological details from their mother tongue. This phonological information comprises the actual speech sounds (phonemes), the rules according to which these sounds are combined (phonotactic knowledge), and the prosodic patterns that help to structure the language input into units, such as information about a word's stress pattern and a sentence's intonational contour.

3.1.1 *Discrimination of phoneme parameters*

In addition to the perception of longer phonological units such as words and sentences, infants have to tackle the basic speech sounds of their mother tongue. The smallest sound units of a language, phonemes, are contrastive from each other, though functionally equivalent. In a given language, a circumscribed set of about 40 phonemes can be combined in different ways to form unique words. A word's meaning will change when one of its component phonemes is exchanged with another, as from *cat* to *pat*. Rather than being discrete phonetically, phonemes are really categories of speech sounds, independent of actual acoustic differences caused by different speakers or by neighboring phonemes, as in the *c* in *cat* versus *cut*.

An ERP study in our laboratory investigated infants' ability at the age of 2 months to discriminate between different vowel lengths in phonemes (Friederici, Friedrich, & Weber, 2002). Infants were presented with two syllables of different duration, /ba:/ vs. /ba/ (see Figure 2 (a)). In order to gain an electrophysiologically measurable response the *mismatch paradigm* was used. In this paradigm, also dubbed *passive oddball paradigm*, two classes of stimuli are repeatedly presented with one stimulus occurring relatively frequently (standard) and the other one relatively rarely (deviant or oddball). As already mentioned, the Mismatch Negativity (MMN) component is a pre-attentive electrophysiological response that is elicited by any discriminable change in repetitive auditory stimulation (Näätänen 1990). In other words, the mismatch response in the ERP is the result of the brain's automatic detection of the deviant among the standards. In our experiment, two separate experimental runs tested the long syllable /ba:/ as deviant in a stream of short syllable /ba/ standards, and short /ba/ as deviant in a stream of long /ba:/ standards.

Figure 2 (b) displays the ERP data and the difference wave obtained from the subtraction of the brain response to the standard stimuli from the one to the deviant stimuli. The critical comparison is between responses to a stimulus as deviant or as standard, independent of its actual physical features. The difference wave shows a positivity with a frontal maximum at around 500 ms post-syllable onset. However, this positivity was only present for the deviancy detection of the long syllable among the short syllables (left panel of Figure 2). In adults, the same experimental setting evoked a different ERP component consisting of a pronounced negative deflection at around 200 ms post-stimulus onset in the difference wave, the typical MMN response to acoustically deviating stimuli; a comparison of the infant and adult ERP data is provided in Figure 3. Interestingly, in infants the response varied dependent on their state of alertness during the experiment. Children who were in quiet sleep during the experiment showed only a positivity, while children who were awake showed an adult-like MMN in addition to the positivity. From the data it follows that infants as young as 2 months of age are



Acoustic Parameters	0 - 30 ms
1st formant	776 Hz
2nd formant	1436 Hz
3rd formant	2735 Hz
4th formant	4086 Hz
Pitch	219 Hz
Pitch min	216 Hz
Pitch max	220 Hz

Figure 2. Syllable discrimination: (a) The figure displays the acoustic parameters of the short syllable /ba/ (202 ms) and the long syllable /ba:/ (341 ms). In a passive auditory oddball paradigm, both syllables are repeatedly presented in the specified frequency (standard 5/6, deviant 1/6). The critical comparison concerns each syllable in its role as standard in one block and deviant in the other. (b) ERP data and difference waves (deviant-standard) of 2-month-olds in response to the long syllable /ba:/ and the short syllable /ba/ (modified from Friederici et al. 2002).

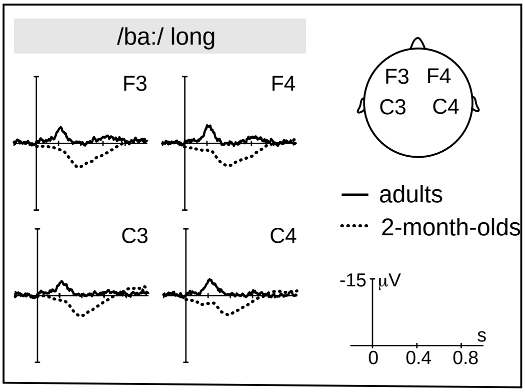


Figure 3. Syllable discrimination: Difference waves (deviant-standard) of adults and 2-month-olds in response to the long syllable /ba:/ in an auditory oddball paradigm (modified from Friederici et al. 2002).

able to discriminate long syllables from short syllables and that they display a positivity in the ERP as mismatch response (MMR). Here, they more easily discriminate a long syllable in a stream of short syllables than vice versa, which can be explained by the greater perceptual saliency of a larger element in the context of smaller elements than vice versa (for more detail see Friedrich, Weber, & Friederici 2004).

So far, several ERP studies have studied phoneme discrimination in infants testing their ability to detect changes in consonant articulation (Dehaene-Lambertz & Dehaene 1994), consonant duration (Kushnerenko et al. 2001; Leppänen et al. 2002), vowel duration (Leppänen, Pikho, Eklund, & Lyytinen 1999; Pihko et al. 1999), and vowel type (Ceponiene, Lepistö, Alku, & Näätänen 2003; Cheour et al. 1998). In these studies, the MMR appeared either as a positive or a negative deflection in the ERP. For instance, Kushnerenko et al. (2001) presented sleeping newborns with fricatives of different duration and observed negative MMRs. Comparatively, Leppänen et al. (1999) and Pihko et al. (1999) investigated sleeping newborns by presenting phonemes with different vowel length and reported positive MMRs. There is an ongoing discussion about the nature of these different ERP responses to auditory change detection. Our studies, for instance, showed that the outcome of these ERP responses are affected by at least two factors: the infants' state of alertness (awake or asleep) and the particular data filtering technique (see Weber, Hahne, Friedrich, & Friederici 2004). Furthermore, the choice of stimulus (discrimination difficulty or saliency) seems to have an impact on the discrimination response (Morr, Shafer, Kreuzer, & Kurtzberg 2002). Also, the transition from a positive to a negative MMR can be shown to be an effect of advancing maturation (see Kushnerenko et al. 2002; Morr et al. 2002; Trainor et al. 2003; for a discussion of a possible overlap between the two components, see Morr et al. 2002). Despite the differences in the appearance of the detection of phoneme changes in the ERP, the combined data suggest that the infants' ability to automatically discriminate between different phonemes is present from early on. Here, ERPs might be used to mark differences in the infants' maturation state that are beyond the scope of observation.

In addition to investigating syllable discrimination in infants with normal language development, we pursued in our research the hypothesis that a major underlying cause of *Specific Language Impairment (SLI)* is a deficiency in the processing of phonological/prosodic information (see Friedrich, Weber, & Friederici 2004). SLI is defined as impairment in the expressive language domain in the presence of otherwise normal development (see Leonard 1998). Consequently, apart from their language deficits these children show normal intelligence and do not have any neurological, sensory, or motor problems. If the hypothesis of an impaired processing of phonological/prosodic information in SLI holds, infants at risk for SLI might be deficient in detecting duration changes in phonemes at the age of 2 months.

To address this question, another study carried out by our laboratory tested the ability of children with a risk for SLI to discriminate long from short syllables with the same MMN procedure as used before (Friedrich et al. 2004). Children were assigned to one of two groups, being at risk or not at risk for later developing SLI, based on family history. In Figure 4, the difference waves for both groups, 2-month-olds with and without risk for SLI, show a positive deflection. However, in the risk group this positive wave reaches its maximum later than the one in the no-risk group. This latency difference points to a slower speed in information processing in the at-risk infants.

Thus, infants at risk for SLI differ from those with no risk at the age of 2 months in their perceptual ERP parameters for duration discrimination. This is in line with recent ERP studies that investigated early differences in phoneme processing between infants with and without a family risk of dyslexia. Specifically, Leppänen et al. (2002) studied ERP responses to consonant duration changes and observed that 6-month-olds at-risk differed from controls in their initial responsiveness to sounds as well as their ability to discern sound changes. Guttorm et al. (2005) reported for at-risk newborns longer lasting positive ERP responses to consonant-vowel changes that were correlated with poorer receptive language skills measured at 2.5 years. Although at this point it is not clear yet whether the children of our study will develop SLI at a later age, the present data suggest that a delayed auditory change processing, possibly caused by weaker memory traces in these children, might be one of the potential factors (for more detail see Friedrich

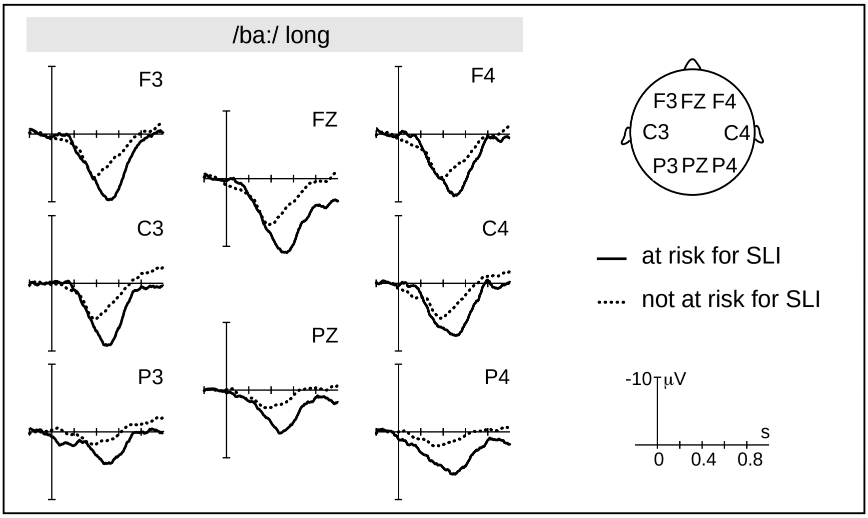


Figure 4. Syllable discrimination: Difference waves (deviant-standard) of 2-month-olds at risk and not at risk for SLI (based on family risk) in response to the long syllable /ba:/ in an auditory oddball paradigm (modified from Friedrich, Weber, & Friederici 2004).

et al. 2004). Consequently, if these children already have difficulties at early language learning stages in detecting phonological/prosodic cues, they might be delayed or impaired in utilizing this information at later stages of lexical and syntactic learning. Tallal et al. (1996) suggest extensive training with artificially slowed speech for children with SLI to overcome abnormal perceptual learning present at early developmental stages. Since the mentioned studies demonstrate that the ERP method is able to differentiate between groups of children at risk and not at risk for later language problems, specific electrophysiological parameters, such as the latency of the mismatch response, might be developed as a diagnostic tool for very early identification of children who would benefit from early intervention.

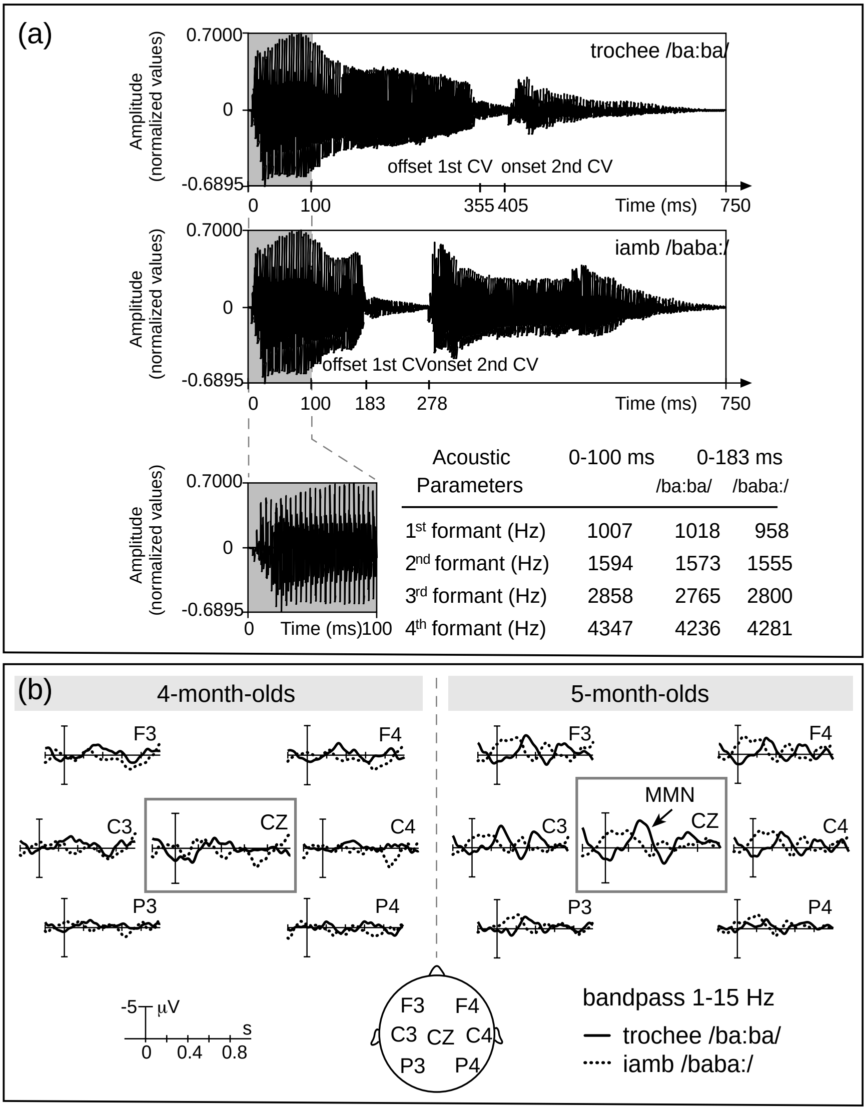
3.1.2 *Discrimination of stress patterns in words*

Another important phonological feature that infants have to discover and make use of during language acquisition is the rule according to which stress is applied to multi-syllabic words. In English and German, for instance, the stress of two-syllable words is realized on the first syllable in more than 90% of the cases, whereas in French the reverse pattern applies. In language acquisition, the detection of words in the speech input, that is, the identification of word boundaries defined by word onset and word offset, is facilitated by information about word stress. Specifically, in languages like English or German, with stress on the first syllable of two-syllable words, stress information certainly aids the identification of word onsets. In order to investigate by means of ERP at what age infants start to show the ability to discriminate word stress, we again applied the MMN paradigm described above.

In this ERP study, 4- and 5-month-old German infants were tested on their ability to discriminate bisyllabic pseudowords stressed on the first syllable from bisyllabic pseudowords stressed on the second syllable (Weber et al. 2004). Since word stress is defined by a number of acoustic parameters with syllable length as the most prominent one, we used pseudowords that differed in the length of their first and second syllables. Namely, we used the trochee /ba:ba/ stressed on the first syllable, and the iamb /baba:/ stressed on the second (see Figure 5 (a)). In a passive auditory oddball paradigm, infants were repeatedly presented with deviant /ba:ba/ stimuli among standard /baba:/ stimuli and vice versa.

Figure 5 (b) displays the difference waves for both age groups. A significant negative deflection is only present in the 5-month-old group, for whom only the trochee evokes a negative MMR. This indicates that by the age of 5 months infants are able to discriminate word stress in bisyllabic words, whereas 4-month-olds are not. The discrimination response is evident in the negative MMR to the trochaic pattern, which is the predominant pattern in German. The negative MMR, in contrast to the positive MMR observed in our 2-month-olds (Friederici et al. 2002), might be attributable to both the infants advanced maturation state as well as the lower processing demands of word stress discrimination (or the higher saliency of the stress differences) than differences in vowel duration (see Kushnerenko et al. 2002; Morr et al. 2002).

Behavioral studies have demonstrated that the ability to discriminate stress patterns in bisyllabic words is not present in 6-month-old infants, but has emerged by the age of 9 months (e.g., Jusczyk, Cutler, & Redanz 1993). In support of the prosodic bootstrapping approach, Nazzi, Iakimova, Bertoncini, Frédonie, and Alcantara (2006) describe developmental effects for 8- to 16-month-olds for the detection of syllables in fluent speech before two-syllable words are derived as one unit. While no segmentation effect was found for 8-month-olds, 12-month-olds segmented individual syllables from the speech stream, with more ease in segmenting the second syllable, which is consistent with the rhythmic features of their native language French. Interestingly, by the age of 16 months children segmented bisyllabic words as whole units from the speech input. Although the results of behavioral and ERP studies are not directly comparable, it seems that before a discrimination reaction can be observed at a behavioral level, there is evidence for a discrimination response in the measured brain activity. This is supported by the fact that our study showed a stress pattern discrimination response in the ERP even in 5-month-olds (Weber et al. 2004). Similarly, Kooijman and colleagues (Kooijman, Hagoort, & Cutler 2005, 2006) found word recogni-



Acoustic Parameters	0-100 ms	0-183 ms /ba:ba/	0-183 ms /baba:/
1st formant (Hz)	1007	1018	958
2nd formant (Hz)	1594	1573	1555
3rd formant (Hz)	2858	2765	2800
4th formant (Hz)	4347	4236	4281

Figure 5. Stress pattern discrimination: (a) The figure displays the acoustic parameters of the bisyllabic pseudowords /ba:ba/ (trochee) and /baba:/ (iamb). In a passive auditory oddball paradigm, both pseudowords are repeatedly presented in the specified frequency (standard 5/6, deviant 1/6). The critical comparison concerns each pseudoword in its role as standard in one block and deviant in the other. (b) Difference waves (deviant-standard) of 4- and 5-month-olds in response to the trochaic and the iambic stress pattern (modified from Weber et al. 2004).

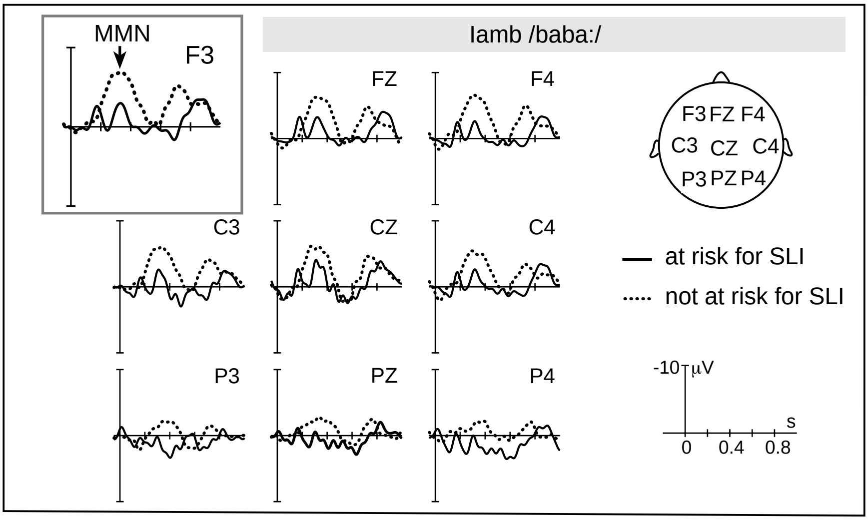


Figure 6. Stress pattern discrimination: Difference waves (deviant-standard) of 5-month-olds at risk and not at risk for SLI (based on word production at two years) in response to the trochaic stress pattern /baba:/ in an auditory oddball paradigm (modified from Weber et al. 2004).

tion responses in the ERP to previously familiarized words in a study testing 10-month-olds and even 7-month-olds. However, behavioral studies also find word segmentation effects around the age of 7–8 months when the words' stress patterns follow the rhythmic features of the infants' native language (e.g., Houston, Santelmann, & Jusczyk 2004; Jusczyk, Houston, & Newsome 1999). Together, these findings emphasize the importance of stress cues for word segmentation from fluent speech.

To follow up on the finding that 5-month-old infants are able to discriminate differently stressed pseudowords, we aimed to test whether infants at risk for SLI already show an impaired stress pattern discrimination at the age of 5 months. Another study carried out in our laboratory investigated the ERP responses of infants at risk for SLI by using the same MMN paradigm with bisyllabic pseudowords (see Weber et al. 2004).

In the current experiment, infants were retrospectively grouped into infants being at risk or not at risk for later SLI based on their word production performance at the age of 24 months (Weber, Hahne, Friedrich, & Friederici 2005). Children who at 24 months have very low word production (as assessed by the ELFRA-2 measure; Grimm & Doil 2000) display at 5 months a reduced MMR amplitude to the trochaic pattern compared to their age-matched controls, as can be observed in the difference waves in Figure 6. It follows that infants with risk for SLI, as determined by a deficit in word production at the age of 24 months, seem to have an impaired discrimination ability for the language-specific trochaic pattern already at the age of 5 months. This result gives rise to the notion that the processing and acquisition of phonological information in early infancy might be crucial for later normal language development. In this case, the ERP mismatch response could serve as an early identification of SLI.

The combined results of the studies on phoneme and stress pattern discrimination provide evidence that speech segmentation in early infancy heavily depends on phonological/prosodic cues and that these cues are likely contributors to lexical acquisition. In their behavioral experiments, Nazzi, Dilley, Jusczyk, Shattuck-Hufnagel, and Jusczyk (2005) demonstrated that indeed both stress pattern and type of initial phoneme influence word segmentation from fluent speech with a preference for the predominant patterns of the infants' native language.

3.1.3 *Phonotactic knowledge*

For successful language learning, infants need not only to be able to recognize and discriminate phonemes but also to know something about the rules according to which these phonemes are combined. As infants become more familiar with the actual sounds of their native language, they also gain probabilistic knowledge about particular phonotactic rules. Phonotactic information refers to the specific rules that define how phonemes may or may not be combined to form words in a given language. This also includes information about which phonemes or phoneme combinations can legally appear at a word's onset or offset. If infants acquire this kind of information from early on, it can support the detection of lexical units in the speech stream and thus aid the learning of new words.

Another study carried out in our laboratory addressed the question whether infants between 12 and 19 months of age possess phonotactic knowledge and, if so, whether they are able to apply this knowledge in lexical processing. The N400 component served as an electrophysiological measure to tease apart infants' responses to pseudowords that follow the phonotactic rules of their native language and nonsense words that do not (Friedrich & Friederici 2005a). The N400 component is known to indicate lexical-semantic processing by marking the effort to integrate an event into its semantic context (for more detail see Holcomb 1993). The N400 amplitude is more pronounced the more semantically unfamiliar, unexpected, or non-matching an event is given the current semantic context or the semantic knowledge in long-term memory. ERP studies examining lexical processing in adults have shown that the N400 amplitude is larger for pseudowords, i.e., phonotactically legal but non-existent in the lexicon, than to real words. However, nonwords, i.e., phonotactically illegal words, do not elicit an N400 response (e.g., Bentin, Mouchetant-Rostaing, Giard, Echallier, & Pernier 1999; Holcomb 1993; Nobre & McCarthy 1994). Pseudowords seem to trigger search processes in

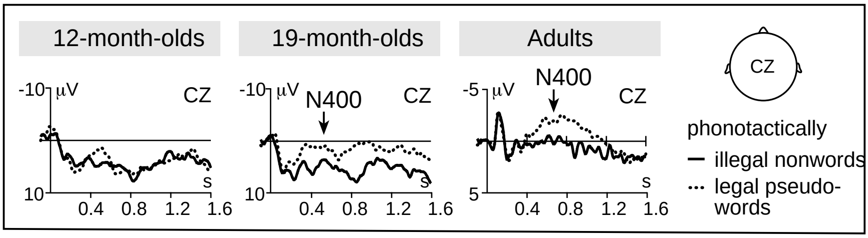


Figure 7. Phonotactic knowledge: ERP data of 12-month-olds, 19-month-olds, and adults in response to pseudowords (phonotactically legal) and nonwords (phonotactically illegal) in a picture-word paradigm (modified from Friedrich & Friederici 2005a).

the lexicon for possible entries, but this search fails as pseudowords are not part of the lexicon. Nonwords, however, do not initiate a similar search response. Apparently, nonwords are not even treated as likely entries of the lexicon and thereby possible referents for meaning, since they already violate the phonotactic rules of the language.

We investigated children's ERP responses to phonotactically legal pseudowords and phonotactically illegal nonwords to determine whether 12- and 19-month-olds already have some phonotactic knowledge (Friedrich & Friederici 2005a). In a p*icture-word paradigm* children were presented with simple colored pictures while simultaneously listening to basic level words spoken slowly. These words either correctly labeled the picture content or were pseudowords or nonwords. Assuming that the picture content initiates lexical-semantic priming, semantic integration difficulties (reflected in an enhanced N400 amplitude) should occur for pseudowords and nonwords, since these nonsense words do not semantically match the picture content. However, if infants at the age of 12 and 19 months posses some phonotactic knowledge, the N400 semantic violation response should only appear for pseudowords, which follow the language-specific phonotactic rules but not for nonwords, which violate them and are therefore not treated as likely referents for picture labels.

As can be seen in Figure 7, the ERP responses of 19-month-olds are quite similar to the ones observed in adults. Both groups show a more negative response to phonotactically legal pseudowords than to phonotactically illegal nonwords. Adults show the typical N400 response, starting at around 400 ms after stimulus onset, most pronounced at central and parietal electrode sites. In 19-month-olds, the more negative deflections to pseudowords also start at around 400 ms post-stimulus but are sustained longer than in adults, suggesting an N400-like response in these children. In contrast, 12-month-olds do not show differential ERP responses to pseudowords and nonwords comparable to those of adults and 19-month-olds. The 12-month-olds do show more negativity with pseudowords, mainly at left lateral frontal sites between 800–900 ms after stimulus onset, which suggests some kind of facilitated processing of phonotactically legal pseudowords. However, this facilitation does not seem to be based on lexical-semantic processing, as marked by an N400 response, but is likely to reflect favored acoustic-phonological processing for the familiar phoneme sequences in phonotactically legal pseudowords (for a more detailed discussion of the effects see Friedrich & Friederici 2005a).

The results of this ERP study on processing of phonotactic information are in line with behavioral studies that show that 9-month-old infants already know some phonotactic rules of their native language (Friederici & Wessels 1993; for a review see Jusczyk 1997). Here, infants were able to use their knowledge about phonotactic rules regarding word boundaries to segment word-like units when pseudowords were presented either in isolation or in context.

In summary, we can conclude that 19-month-olds possess some phonotactic knowledge (indicated by an N400) and therefore treat as potential words phonotactically legal pseudowords but not phonotactically illegal nonwords. Thus, phonotactically illegal nonwords are from very early on excluded from further word learning. 12-month-old infants seem not to have established this kind of phonotactic knowledge yet but instead show a phonological familiarity effect to phonotactically legal pseudowords. Taken together with the observation from behavioral studies that 9-month-old infants have some initial phonotactic knowledge one may conclude that this kind of knowledge advances and is applied in lexical processing (and in this case marked by the occurrence of an N400) only a few months later in the child's development. This assumption is supported by the finding that the mechanisms that underlie the N400 response mature between 12 and 14 months (see Friedrich & Friederici 2005a, 2005b).

3.1.4 *Processing of intonational phrase boundaries*

The studies just discussed demonstrate that lexical learning is based on phonological knowledge at the segmental level. Let us now turn to the suprasegmental level, such as the intonational contours in phrases and sentences. The location of phonological and intonational phrase boundaries has an effect on lexical learning, since it marks the beginning or the end of lexical entities and consequently aids the detection of words. Furthermore, when infants start to tease apart the syntactic units and to discover the syntactic rules of their native language, prosodic information at the sentence level plays a central role. The identification of syntactic phrases is eased by prosodic information since each intonational phrase boundary is a syntactic phrase boundary, although the reverse relation does not hold in some cases (for a review on the relationship between prosody and syntax see Cutler, Dahan, & van Donselaar 1997). The detection and processing of intonational phrase boundaries thus could provide an easy entry into the lexicon and the syntax of a given language.

Intonational phrase (IPh) boundaries are defined by several characteristics. First, the last syllable of an IPh contains a change in pitch, marking a low or a high boundary tone. Second, this last syllable is lengthened as compared to syllables within the phrase. Third, there is often a pause following the IPh boundary. Behaviorally, it has been shown that adult listeners make use of IPh boundaries in the interpretation of spoken utterances (see Cutler, Dahan, & van Donselaar 1997). In ERP studies with adults, the offset of intonational phrases is associated with a positive deflection with a centro-parietal distribution (Pannekamp, Toepel, Alter, Hahne, & Friederici 2005; Steinhauer, Alter, & Friederici 1999). This component is named *Closure Positive Shift (CPS)* since it is interpreted as an indicator for the closure of prosodic phrases by intonational phrase boundaries. In order to gain further insights into the role of prosodic information in language acquisition, we have investigated the neurophysiological basis of the perception of sentence-level prosodic cues in infants.

In an ERP study examining the ability of 8-month-old infants to identify IPh boundaries, Pannekamp, Weber, and Friederici (2006) presented infants with sentences of two prosodically correct sentence types, one consisting of two IPhs with one sentence-internal IPh boundary (see example 1) and the other one containing three IPhs with two sentence-internal IPh boundaries (see example (2)). The two sentence conditions resulted from the different syntactic structure of the sentences. The use of intransitive versus transitive verbs leads to late closure versus early closure, respectively. The different syntactic structures consequently result in different prosodic realizations with IPh boundaries at different sentence positions.

(1) Lena verspricht Mama zu flitzen und Getränke zu kaufen.
[Lena promises Mama to run]$_{\text{IPh}}$ [and drinks to buy]

(2) Lena verspricht Mama zu helfen und Getränke zu kaufen.
[Lena promises]$_{\text{IPh}}$ [Mama to help]$_{\text{IPh}}$ [and drinks to buy]

The ERP results in response to the two different sentence types for both the infant study (Pannekamp et al. 2006) and the adult study (Pannekamp et al. 2005) are given in Figure 8. In both groups, the occurrence of IPh boundaries over the course of the sentence is followed by a positive shift in the ERP. Thus, for sentences with one IPh boundary one corresponding CPS was observed, and for sentences with two IPh boundaries, two CPS. In adults, the positive shifts start with an approximate latency of 500 ms to their corresponding IPh boundaries, whereas in infants the positive waves start about 1000 ms after the IPh offset. This developmental latency shift has already been reported in prior infant ERP

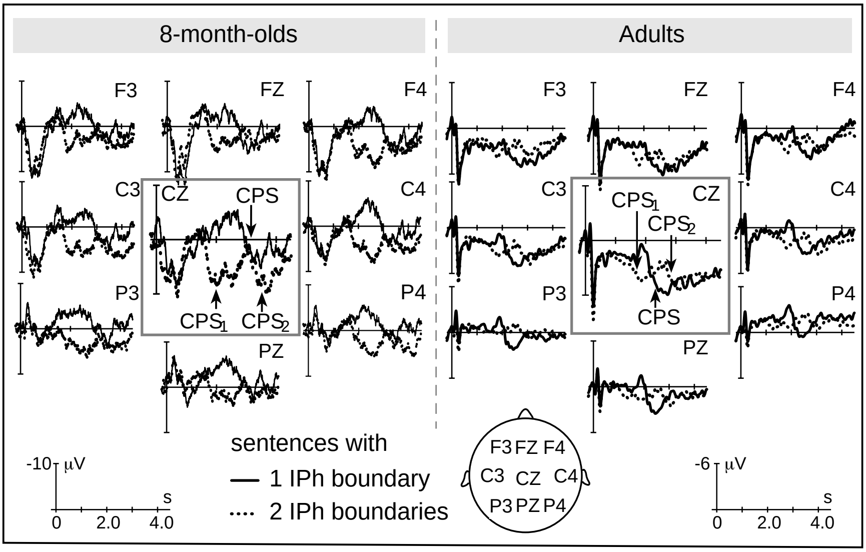


Figure 8. Processing of intonational phrase boundaries: ERP data of 8-month-olds and adults with positive shifts (CPS) in correlation to the IPh boundaries in each sentence condition (modified from Pannekamp et al. 2005, 2006).

studies on semantic and syntactic processes and points to slower information processing in infants and children (see for instance Hahne, Eckstein, & Friederici 2004; Oberecker, Friedrich, & Friederici 2005). The CPS latency to each of the IPh boundaries further indicates that this ERP component is not a mere reaction to the onset of acoustic differences (lower-level processing) but an index for the underlying linguistic process of the perception of intonational phrases (higher-level processing). Since the infant ERP data clearly show the occurrence of the CPS component in response to each IPh boundary, the current study indicates in electrophysiological terms that infants as young as 8 months possess the ability to process prosodic cues at the sentence level. This finding supports the notion that the neurophysiological basis of prosodic processes that are crucial for speech segmentation is established early during language development.

The results of this ERP investigation are in line with behavioral studies that show that infants at the age of around 9 months are able to detect major syntactic phrases in the speech input by relying on prosodic information (Hirsh-Pasek et al. 1987; Jusczyk et al. 1992; for a review see Jusczyk 1997). Recently, Soderstrom and colleagues (Soderstrom, Nelson, & Jusczyk 2005; Soderstrom, Seidl, Nelson, & Jusczyk 2003) have shown that even 6-month-olds are sensitive to phrase-level prosodic cues and that the infants indeed utilize this information to identify syntactic units in the speech stream. Similarly, in lexical learning, behavioral studies have shown that 10-month-old infants use phonological phrase boundary information to constrain lexical access (Christophe, Gout, Peperkamp, & Morgan 2003; Gout, Christophe, & Morgan 2004). Taken together, these studies demonstrate that infants are not only able to perceive phrase-level prosodic cues in early infancy but that they actually use this information to segment fluent speech into lexical and syntactic units. Future ERP studies may provide additional information about the on-line brain mechanisms when infants apply prosodic information to structure the speech input. While behavioral studies deliver the information that infants perceive syntactic units on the basis of prosodic cues, ERP studies may show how and when these cues are utilized to disentangle the speech input.

3.2 Processing of lexical-semantic information

Another challenge that infants have to face in language learning is the acquisition of the meanings of words. Infants successively build up their lexicon and develop the ability to map words onto conceptual representations of objects or events and vice versa. From early on infants segment words from fluent speech, but how and when they actually know what a particular word means still needs to be specified. To answer this question and to understand the neurophysiological mechanisms of early word learning, it is important to determine whether the infant's brain works similarly to the adult brain when processing meaningful words in meaningful contexts, such as pictures or sentences.

In the adult ERP research on semantic processing, the N400 component has been established as an indicator for the neural mechanisms of semantic integration of elements into their context (e.g., Holcomb 1993). The N400 is evoked in response to both words and pictures that do not match the expectation build up by previously presented words, sentences, pictures, and picture stories (e.g., Friederici, Pfeifer, & Hahne 1993; West & Holcomb 2002). The amplitude of the N400 is inversely related to the expectation triggered by the semantic context and therefore varies with the effort necessary to integrate a stimulus into a given situation. An N400 priming effect, reflected in a reduced N400 amplitude, thus indicates the effect of semantic priming on the process of semantic integration. Importantly, a reduced amplitude reflects ease in semantic integration due not only to previously presented stimuli, i.e., the current semantic context, but also to the semantic knowledge in long-term memory (for a review see Kutas & Federmeier 2000). Since the N400 has been successfully shown to be correlated to lexical-semantic processing in adults and since this component is a likely indicator for the development of semantic memory, we have utilized the N400 to study the neurophysiological basis and the developmental aspects of lexical-semantic processes in early childhood.

3.2.1 Processing of lexical-semantic information at word level

Several studies in our laboratory have focused on the question of whether the neural mechanisms for semantic processing observed in adults are also present during early language acquisition when infants have acquired only a few words. Furthermore, these studies aim to determine at what point during the infants' development these neural mechanisms mature, by examining infants' abilities during word-level lexical-semantic processing as indexed by the N400 component (Friedrich & Friederici 2004, 2005a, 2005b, 2006). Using a cross-modal picture-word design, the ERP responses of 12-, 14-, and 19-month-old children were recorded to slowly spoken, basic level words. While the children were looking at sequentially presented pictures, they were acoustically presented with words that were either congruous or incongruous to the picture content. If infants are already able to integrate word meaning into semantic context, provided here by the contents of the pictures, an N400 priming effect should occur. The successful comprehension of words should be reflected in a reduced N400 amplitude for words that match their picture, while an enhanced N400 amplitude should be present for non-matching words.

In adults, the described paradigm evokes an N400 in the ERP in response to the incongruous words (see Figure 9; Friedrich & Friederici 2004). This negative deflection is most pronounced in the time range of 400 to 800 ms post-stimulus onset, predominantly at centro-parietal electrode sites, but also extends to frontal sites and lasts up to 1200 ms. As shown in Figure 9, in 19-month-old children an N400-like effect to incongruous words was observed (Friedrich & Friederici 2004). This negativity, with a centro-parietal maximum, starts at about 400 ms post-stimulus and sustains up to 1400 ms. In children, the effect starts later and lasts longer than in adults, also showing a stronger involvement of the frontal electrode sites. From this it follows that even 19-month-old children show a long-lasting incongruity effect, which points to their ability to process the meaning of a word in its context. The picture content successfully activates the associated semantic knowledge that facilitates or hampers subsequent semantic processing when a word matches or does not match the picture content, respectively.

Another study carried out in our laboratory examined these early word learning abilities in correlation to the children's later expressive language skills (Friedrich & Friederici 2006). As described earlier, ERP data of 19-month-olds were retrospectively grouped based on the children's word and sentence production

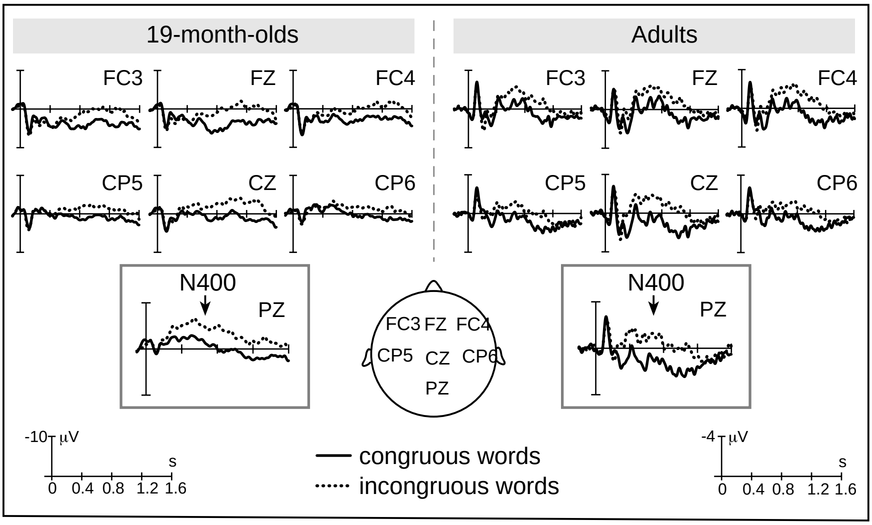


Figure 9. Processing of lexical-semantic information at word level: ERP data of 19-month-olds and adults in response to semantically incongruous and congruous words in a picture-word paradigm (modified from Friedrich & Friederici 2004).

performance at the age of 30 months (German language development test SETK-2; Grimm 2000). According to the children's test results 19-month-olds were classified based on whether they had later age-appropriate language skills or on whether they were at risk for later language problems, that is, based on whether they had poor language production skills at the word or sentence level. Since impaired language production is one of the major features of SLI, these children seem to have an enhanced risk for later occurring SLI (see Leonard 1998; Rescorla, Roberts, & Dahlsgaard 1997; Rescorla, Bascome, Lampard, & Feeny 2001). As displayed in Figure 10, children at risk for SLI do not show the same N400 effect in response to the incongruous words as their peers. Hence, children who show language production deficits at the age of 30 months seem to be impaired in their word processing abilities about one year before they experience problems in their expressive language skills. The results suggest that the emergence of the N400 during the development is related to children's later expressive language skills and that the missing N400 response indicates a delayed semantic development in the children with later language problems. From this it follows that the ERP method may deliver indicators for SLI risk at an earlier developmental stage than behavioral measures. While it is challenging to test children who know only a few words on their expressive language skills, electrophysiological parameters can be obtained from early on.

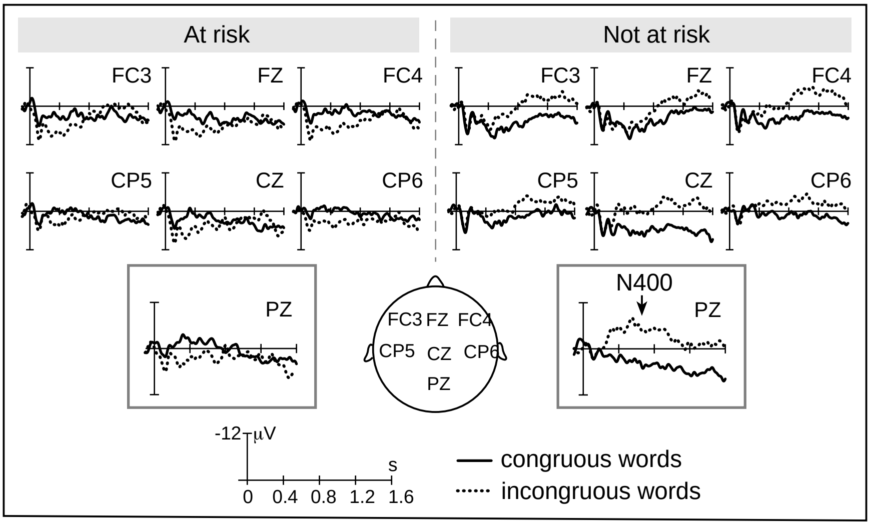


Figure 10. Processing of lexical-semantic information at word level: ERP data of 19-month-olds at risk for SLI (impaired language production skills at 30 months) and not at risk for SLI (normal language production skills at 30 months) in response to semantically incongruous and congruous words in a picture-word paradigm (modified from Friedrich & Friederici 2006).

To pinpoint at what developmental stage the neural processes that underlie the N400 component mature, another study of our laboratory focused on the developmental progression of word processing abilities across different age groups (Friedrich & Friederici 2005a, 2005b). The ERP responses of 12-, 14-, and 19-month-olds to incongruous and congruous picture-word pairs were compared (see Figure 11). Interestingly, the N400-like effect to incongruous words was not only present in 19-month-olds but also in 14-month-olds, although with a broader distribution in the younger age group. In contrast, in 12-month-old infants this effect is not yet established. However, these youngest children show an early negativity in the time range of 100 to 500 ms post-word onset in response to congruous words (Figure 11). This early negative effect at lateral frontal electrode sites was also present in all other age groups. This negativity could be interpreted as an early phonological-lexical priming effect for the congruent picture-word pairs. More specifically, the authors propose a facilitation of the acoustic-phonological processing of the congruous words, which were expected due to lexical-semantic priming by the pictures, whereas the incongruous words were not (for more detail see Friedrich & Friederici 2005a, 2005b). Thus, even 12-month-olds seem to have an initial lexical knowledge that is activated by the picture content and favors the acoustic-phonological processing of expected words. The absence of an N400 in these children,

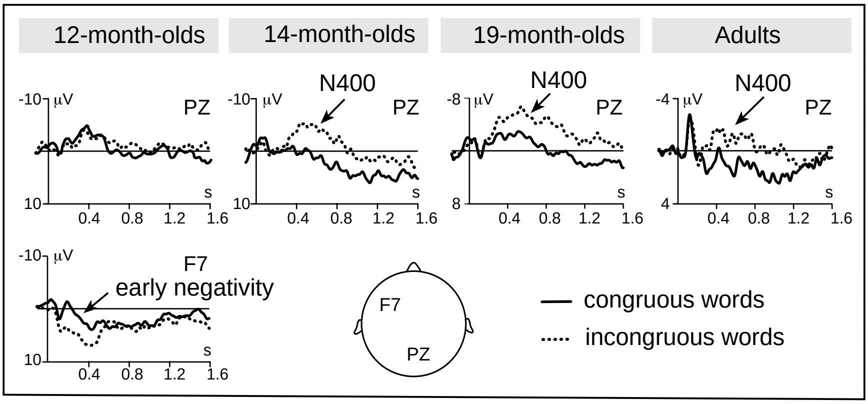


Figure 11. Processing of lexical-semantic information at word level: ERP responses of 12-, 14-, 19-month-olds, and adults in response to semantically incongruous and congruous words in a picture-word paradigm (modified from Friederici 2005).

however, supposes that their semantic memory structures are still too unspecific or unstable to trigger the mechanism underlying the N400 generation.

Similarly to the described results, early negative effects in the ERP have been observed in studies on the processing of known and unknown words. Negative effects in the time range of 200 to 400 ms post-word onset were more pronounced in response to known words than to unknown words in 20- and even 14-month-old children (Mills, Coffey-Corina, & Neville 1993; 1994; Mills, Prat, Zangl, Stager, Neville, & Werker 2004). Interestingly, in 20-month-olds but not yet in 14-month-olds, ERP responses even differentiated between known words and those words, which differed in their initial phonemes, with more negative responses to the former. Also, Thierry, Vihman, and Roberts (2003) found early negative ERP effects in response to familiar words as opposed to unfamiliar words in the time range between 170 to 240 ms post-word onset. The combined results of these studies support our view that the early negative response at lateral sites, also observed in our study for congruous words, reflects a familiarity effect. This is in line with ERP effects in 10-month-olds, where after several presentations of low frequency words a negativity emerges (Kooijman et al. 2005). In general, facilitated phonological processing reflected in the early negativity might be affected by both familiarity on the basis of repeated presentation and existing phonological-lexical representations in the long-term memory.

In summary, in the current ERP studies on semantic processing at the word level, two ERP effects are observed. First, an early negativity in response to congruous words is present in all age groups, even in the 12-month-olds. We interpret this early negative effect as reflecting facilitation of phonological processing by lexical priming. And second, the later centro-parietal negativity in the ERP for incongruous words is present in 14- and 19-month-olds, an effect referred to as infant N400. In contrast to their peers, 19-month-olds at risk for later language impairment do not show the expected effect of semantic processing reflected in the N400. Since even 14-month-olds show an N400-like component, this implies that at-risk children are delayed in their semantic development for at least half a year. The occurrence of a phonological-lexical priming effect in all age groups indicates that not only 14- and 19-month-olds but also 12-month-olds already create lexical expectations from picture contents, revealing that they already possess some lexical-semantic knowledge. However, infants at that age do not yet display an N400 semantic incongruity effect that is present in 14-month-olds. From that we assume that the neural mechanisms of the N400 mature between 12 and 14 months of age.

The described ERP studies show that the N400 can be utilized as an indicator for word comprehension in a given semantic context even in early childhood. The result that the N400 is present in the second year of life implies that this ERP component is a useful tool to further investigate higher-level semantic processing in very young children. In addition to studying the processing of single semantic units, the question arises when children start to process semantic relations between these units, for instance in sentential context.

3.2.2 *Processing of sentence-level lexical-semantic information*

In their language learning environment, infants are usually not presented with words in isolation but in the context of sentences. Sentence comprehension is a complex process that requires the listener to maintain sequentially arriving information in working memory and to semantically integrate this information over time. To understand the meaning of a sentence, infants furthermore need to have some semantic knowledge about verbs and nouns as well as about their respective relation. However, it is not clear whether children at the age of only one or two years are able to process word meaning and semantic relations in sentential context similarly to adults. In order to investigate children's processing of meaningful words in sentence context, the *semantic violation paradigm* can be applied. In this paradigm, semantically correct and incorrect sentences are presented, such as *The king was murdered* and *The honey was murdered*, respectively (Friederici et al. 1993; Hahne & Friederici 2002). The semantic violation paradigm again utilizes the N400 component as an index of semantic integration abilities, with larger N400 amplitudes for higher integration efforts of semantically inappropriate words into their context. In sentence comprehension, the listener sequentially builds up a semantic expectation about the likely ending for a sentence, with a restricted number of possible word candidates. For *The honey was murdered*, this

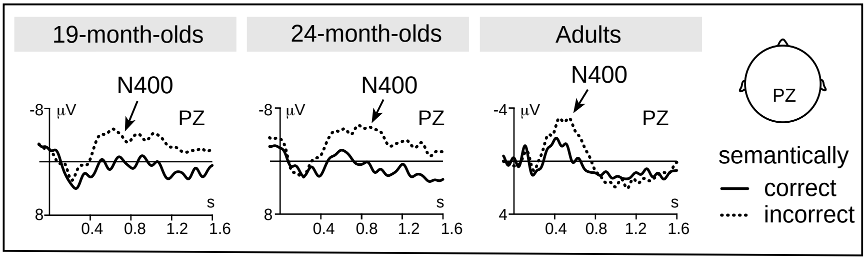


Figure 12. Processing of lexical-semantic information at sentence level: ERP data of 19-month-olds, 24-month-olds, and adults in response to the sentence endings of semantically incorrect and correct sentences in a semantic violation paradigm (modified from Friedrich & Friederici 2005c).

expectation is violated when the verb at the end of the sentence (*murdered*) does not semantically meet the meaning that was set up by the noun in the beginning of the sentence *(honey).* In the ERP responses to those incorrect sentences an N400 occurs for the semantically unexpected sentence endings, as has been demonstrated in adult studies (Friederici et al. 1993; Hahne & Friederici 2002).

In an infant study conducted in our laboratory, the ERP responses to semantically correct and incorrect sentences were analyzed in 19- and 24-month-old children (Friedrich & Friederici 2005c). Both sentence types followed a simple subject-verb-object structure. Semantically incorrect sentences contained objects that violated the selection restrictions of the preceding verb, as in the sentence *The cat drinks the ball* as opposed to *The child rolls the ball.* If children are able to process word meaning in sentence context and already possess some knowledge about the selection restrictions of particular verbs, an N400 to the unexpected sentence endings should be observable in the ERP for incorrect sentences.

In both groups of children, the sentence endings of semantically incorrect sentences but not of correct sentences evoked N400-like effects in the ERP, with a maximum at centro-parietal electrode sites (Figure 12). In comparison to the adult data, the negativities in children start at about the same time, i.e., at around 400 ms post-word onset, but last longer. This suggests that semantically unexpected nouns that violate the selection restrictions of the preceding verb also initiate semantic integration processes in children but that these integration efforts are maintained longer than in adults. Despite these processing differences, the current data indicate that children at the age of 24 and even 19 months are able to process semantic relations between words in a sentence quite similarly to adults.

Other electrophysiological studies have focused on the processing of lexical-semantic information at sentence level in older children, namely 5- to 15-year-olds (Hahne et al. 2004; Holcomb, Coffey, & Neville 1992). These studies reported

N400-like responses to semantically incorrect sentences for all age groups. Recently, Silva-Pereyra and colleagues investigated sentential semantic processing abilities in preschoolers and in 30-month-old children (Silva-Pereyra, Klarman, Lin, & Kuhl 2005a; Silva-Pereyra, Rivera-Gaxiola, & Kuhl 2005b). In the preschoolers (3- and 4-year-olds), sentence endings that semantically violated the preceding sentence phrases evoked several anteriorly distributed negative peaks. Similarly, in 30-month-olds, an anterior negativity between 500–800 ms post-word-onset was observed in response to semantically anomalous sentences. The distribution of these negativities did not match the usual centro-parietal maximum of the N400. Therefore, it is a question whether these negative ERP responses indeed reflect semantic integration processes. Here, it would have been useful to have adult data from the same experimental paradigm as a baseline to determine whether the used paradigm evokes an N400 component. Nevertheless, these studies show differential responses to semantically incorrect and correct sentences in young children. Despite the different effects reported in the ERP studies on sentential semantic processing, the results of our ERP study suggest that semantic processes at sentence level, as reflected by an N400-like response, are present at the end of children's second year of life.

3.3 Processing of syntactic information

The sole combination of phonological features and semantic units does not yet create meaningful language. In fact, a well-defined rule system is necessary to relate the elements of a sentence together in an organized manner, thereby giving the sentence its structure. The composition of elements is regulated by the syntactic rules of a particular language. These rules are important in both language comprehension and production, since they enable speakers and listeners to communicate who does what to whom. Thus, syntactic rules define the grammatical relations of words and phrases in a sentence. The analysis of syntactic relations between and within phrases is a complicated and sophisticated process, yet children have acquired the basic syntactic rules of their native language by the end of their third year of life. However, so far not much is known about the underlying neurophysiological mechanisms of the particular stages in syntax acquisition.

The ERP method permits investigating the neural correlates of syntactic information processing by applying the *syntactic violation paradigm*. Here, syntactically correct and syntactically incorrect sentences are presented. Syntactic violations are for example realized in morphosyntactic, phrase structure, or tense violations. In adult ERP studies, two components have been observed in response to syntactically incorrect sentences containing phrase structure violations. First, the ELAN, an early anterior negativity, is interpreted to reflect highly automatic phrase-structure building processes (Friederici et al. 1993; Hahne & Friederici 1999). Second, the P600, a later occurring centro-parietal positivity, is taken to reflect processes of syntactic integration (Kaan et al. 2000) and controlled processes of syntactic reanalysis and repair (Osterhout & Holcomb 1993; Friederici, Hahne, & Mecklinger 1996). This biphasic ERP pattern in response to phrase structure violations has been found for both passive as well as active sentence constructions (Friederici et al. 1993; Hahne & Friederici 1999; Hahne et al. 2004; Rossi, Gugler, Hahne, & Friederici 2005).

3.3.1 *The detection of phrase structure violations*

In sentence comprehension, children are confronted with the challenge of having to detect the underlying syntactic structure in the speech input by identifying structural units, such as words and phrases. Here, the relations between and within phrases are essential to the understanding of and learning about the roles of syntactic elements that define, for example, the subject and the object of an action. An approach to investigate the question of at what age children are able to process phrase structure information is to test whether children of different age groups are sensitive to syntactic errors in the form of phrase structure violations. Similarly to the adult studies, the use of the syntactic violation paradigm in children ERP studies, can deliver electrophysiological indicators for the detection of syntactic errors within phrases.

ERP studies in our laboratory have investigated at what age children process phrase structure violations and therefore show the syntax-related ERP components ELAN and P600 that have been observed in adults (Oberecker et al. 2005; Oberecker & Friederici 2006). In these experiments, the EEG signal of 24- and 32-month-old children and adults was recorded, while subjects were listening to syntactically correct, syntactically incorrect, and filler sentences. Syntactic violations were realized by omitting the noun after the preposition in simple active sentences, e.g., **The lion in the __ roars* versus *The lion roars*. Thus, syntactically incorrect sentences contained incomplete prepositional phrases.

In the data analyses, ERP responses to the critical verb of the syntactically incorrect and correct sentences were compared (see Figure 13). The adult data demonstrate the expected biphasic ERP pattern in response to the sentences containing a phrase structure violation. The ERP responses of 32-month-old children show a quite similar pattern to those of adults, although both ERP components appear in later time windows. Interestingly 24-month-old children also demonstrate differential ERP responses to correct and incorrect sentences. However, in this age group only a P600 but no ELAN can be observed.

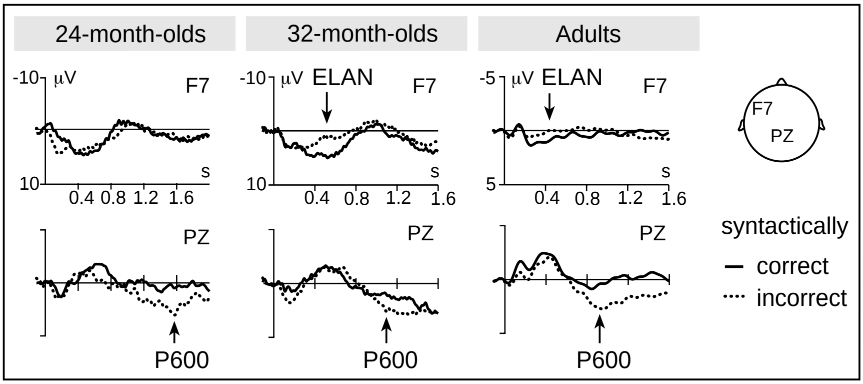


Figure 13. Processing of syntactic violations: ERP data of 24-month-olds, 32-month-olds, and adults in response to syntactically incorrect and correct sentences in a syntactic violation paradigm (modified from Oberecker et al. 2005).

Other neurophysiological studies have addressed the processing of syntactic information in older children. Silva-Pereyra and colleagues (2005a, 2005b) examined the processing of tense violations in active sentences in children between 30 and 48 months. The authors reported a late positivity for the older children and a very late occurring positivity for the 30-month-olds. Furthermore, Hahne et al. (2004) investigated the processing of phrase structure violations in syntactically more complicated passive sentences. The ELAN-P600 pattern was seen in 7-to-13-year-old children; however, 6-year-olds only displayed a late P600 but no ELAN. In comparison to these results, the ERP data of the present studies show for the first time that the child's brain is sensitive to phrase structure violations in active sentences, even at the age of only 24 months. The ERPs of 32-month-olds show both initial processes of structure building (ELAN) as well as late processes of syntactic integration (P600). In contrast, the ERPs of 24-month-old children suggest a developmental change from 2 to 2.5 years since in these children only a P600 component without the expected ELAN occurred. Thus, the data indicate that the two ERP components are somewhat independent and are likely to reflect different processing mechanisms that mature at different stages during the children's language development.

In summary, the combined results indicate that automatic syntactic processes reflected by the ELAN are present earlier for active sentences than for passive sentences during the children's language development. Furthermore, the data suggest that the processes reflected by the P600 are established earlier during the child's development than those reflected by the ELAN. Here, the ERP method, in contrast to behavioral methods, can help to understand and sketch the developmental course of early syntax acquisition, before children can actually produce comparably complex syntactic constructions.

4. Conclusion

The results of the neurophysiological research conducted in our and other laboratories demonstrate that the method of event-related brain potentials is a powerful tool to investigate and monitor early stages of language acquisition. ERPs allow us to describe the underlying neurophysiological mechanisms of the language acquisition process as children develop their receptive language skills. In this context, the described experiments aim to show not only that there are specific ERP indicators of particular language processes in infants and children, but also that these indicators can be used as templates to define the hallmarks of language acquisition.

The described ERP studies in infants and children broadly cover the prosodic, semantic, and syntactic aspects of language acquisition during the first three years of life. We have demonstrated that the ERP method delivers information about the neural correlates of language processes and therefore provides a better understanding of the *how* and *when* of the developmental stages in the language acquisition process. More specifically, we have seen that a particular ERP component that reflects discrimination of phonological features is present even in newborns and can thus be used to examine very early stages of language acquisition. A further component that indicates lexical-semantic processes in adults, the N400, is registered in 14-month-olds, although not yet in 12-month-olds, and can therefore be used to investigate phonotactic knowledge, word recognition and processing of lexical-semantic relations between verbs and their arguments in sentences. For the syntactic domain, an adult-like biphasic ERP pattern, the ELAN-P600, is present in 32-month-old children, but not yet in 24-month-olds, for the processing of structural dependencies within phrases, thus characterizing the developmental progression of the acquisition of syntax.

The comparison of adult and infant ERP data shows that infant ERP components in response to specific processes are sometimes delayed as compared to the adult components, although the components' relative timing matches that of the adult data. Specifically, local phrase structure building precedes lexical-semantic processes and thematic role assignment. Given the latency delay in infant ERP components on the one hand, but the appearance of those components as in the adult data on the other hand, the combined results are in support of a continuity view of language development.

In summary, the ERP method is proving to be a useful research tool in the work with infants and children, especially during early stages of language learning.

The ERP method places virtually no demands on the infants' behavior and delivers an on-line measure of the brain mechanisms underlying the infants' language processing skills. Although we are still far from a detailed outline of the exact steps in the language acquisition process, using the ERP method allows researchers to study language development from very early on and thus provides the possibility to further fill in the gaps in what we know about language development to gain a more fine-grained picture of acquisition and its neurophysiological basis.

5. Future objectives

In the field of developmental neuroscience, researchers aim to achieve a more detailed understanding of the neuronal correlates of language learning. Despite the advantages and the fast advancements of neurophysiological and neuroimaging techniques, this research field considerably benefits from the groundbreaking information provided by behavioral researchers. Thus, effective interdisciplinary communication is essential to scientific progress, since different research areas, using various methods and focusing on different aspects of language learning, all deliver single pieces to the puzzle of language acquisition.

Furthermore, only an interdisciplinary communication between scientists and early child care providers as well as educators ensures an effective use of new findings in the field of developmental cognitive neuroscience. Information might not only help a better grasp of the normal language acquisition process but also deliver new insights into the nature of impaired language acquisition. Importantly, once the ERP method has been further developed so that conclusions can be drawn not only from group data, but also from individual data, certain ERP components might serve as early indicators of an impaired language development. If the ERP method can be utilized as a diagnostic tool, potential language problems can be identified early in the child's development to start intervention.

References

Anderson, A. W., Marois, R., Colson, E. R., Peterson, B. S., Duncan, C. C., Ehrenkranz, R. A., et al. (2001). Neonatal auditory activation detected by functional magnetic resonance imaging. *Magnetic Resonance Imaging, 19*, 1–5.

Bansal, P., Sun, M., & Sclabassi, R. J. (2004). Simulation and extraction of single-trial evoked potentials. *Proceedings of 26th Annual International Conference of the IEEE* 1. *Engineering in Medicine and Biology Society 2004. EMBC 2004,* 200–203.

Bee, H. & Boyd, D. (2007). *The developing child.* Boston MA: Pearson Education.

Berg, P. & Scherg, M. (1994). A multiple source approach to the correction of eye artifacts. *Electroencephalography and Clinical Neurophysiology*, *90,* 229–241.

Bentin, S., Mouchetant-Rostaing, Y., Giard, M. H., Echallier, J. F., & Pernier, J. (1999). ERP manifestations of processing printed words at different psycholinguistic levels: Time course and scalp distribution. *Journal of Cognitive Neuroscience*, *11*(3), 235–260.

Brauer, J. & Friederici, A. D. (2007). Functional neural networks of semantic and syntactic processes in the developing brain. *Journal of Cognitive Neuroscience, 19*(10), 1–15.

Brunia, C. H. M., Möcks, J., van den Berg-Lenssen, M. M. C., Coelho, M., Coles, M. G. H., Elbert, T., et al. (1989). Correcting ocular artifacts in the EEG: A comparison of several methods. *Journal of Psychophysiology*, *3*, 1–50.

Ceponiene, R., Lepistö, P., Alku, H. A., & Näätänen, R. (2003). Event-related potential indices of auditory vowel processing in 3-year-old children. *Clinical Neurophysiology*, *114*(4), 652–661.

Chatrian, G. E., Lettich, E., & Nelson, P. L. (1988). Modified nomenclature for the "10%" electrode system. *Journal of Clinical Neurophysiology, 5*(2), 183–186.

Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., & Näätänen, R. (1998). Development of language-specific phoneme representations in the infant brain. *Nature Neuroscience, 1*, 351–353.

Christophe, A., Gout, A., Peperkamp, S., & Morgan, J. (2003). Discovering words in the continuous speech stream: The role of prosody. *Journal of Phonetics. Special Issue: Temporal integration in the perception of speech*, *31*(3–4), 585–598.

Clark, E. V. (2003). *First language acquisition*. Cambridge: CUP.

Coles, M. G. H. & Rugg, M. D. (1995). Event-related potentials: An introduction. In M. D. Rugg & M. G. H. Coles (Eds.), *Electrophysiology of mind* (pp. 1–26). Oxford: OUP.

Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, *40*(2), 141–201.

Daffner, K. R., Scinto, L. F. M., Calvo, V., Faust, R., Mesulam, M. M., West, W. C., & Holcomb, P. (2000). The influence of stimulus deviance on electrophysiologic and behavioral responses to novel events. *Journal of Cognitive Neuroscience*, *12*(3), 393–406.

Dehaene-Lambertz, G., & Dehaene, S. (1994). Speed and cerebral correlates of syllable discrimination in infants. *Nature, 370,* 292–295.

Dien, J. (1998). Issues in the application of the average reference: Review, critiques, and recommendations. *Behavior Research Methods, Instruments & Computers. Special Issue: Event-related brain potential methodology*, *30*(1), 34–43.

Donchin, E. & Coles, M. G. H. (1988). Is the P300 a manifestation of context updating? *Behavioral and Brain Sciences*, *11*, 357–374.

Edgar, J. C., Stewart, J. L., & Miller, G. A. (2005). Digital filters in ERP research. In T.C. Handy (Ed.), *Event-related potentials: A methods handbook* (pp. 85–113). Cambridge MA: The MIT Press.

Fabiani, M., Gratton, G., & Coles, M. G. H. (2000). Event-related potentials: Methods, theory, and applications. In J. Cacioppo, L. Tassinary & G. Berston (Eds.), *Handbook of psychophysiology* (2nd ed., pp. 52–84). New York NY: Plenum.

Friederici, A. D. (1994). Funktionale Organisation und Reorganisation der Sprache während der Sprachentwicklung: Eine Hypothese. *Neurolinguistik, 8*(1), 41–55.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Science*, *6*, 78–84.

Friederici, A. D. (2004a). The neural basis of syntactic processes. In M. S. Gazzaniga (Ed.), *The Cognitive Neuroscience III* (3rd ed., pp. 789–801). Cambridge MA: The MIT Press.

Friederici, A. D. (2004b). Event-related brain potential studies in language. *Current Neurology and Neuroscience Reports*, *4*, 466–470.

Friederici, A. D. (2005). Neurophysiological markers of early language acquisition: From syllables to sentences. *Trends in Cognitive Sciences*, *9*, 481–488.

Friederici, A. D. & Mecklinger, A. (1996). Syntactic parsing as revealed by brain responses: First-pass and second-pass parsing processes. *Journal of Psycholinguistic Research, 25*, 157–176.

Friederici, A. D. & Wessels, J. M. I. (1993). Phonotactic knowledge of word boundaries and its use in infant speech perception. *Perception and Psychophysics*, *54,* 287–295.

Friederici, A. D., Friedrich, M., & Weber, C. (2002). Neural manifestation of cognitive and precognitive mismatch detection in early infancy. *NeuroReport*, *13*, 1251–1254.

Friederici, A. D., Hahne, A., & Mecklinger, A. (1996). Temporal structure of syntactic parsing: early and late event-related brain potential effects. *Journal of Experimental Psychology: Learning Memory and Cognition, 22*, 1219–1248.

Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: Effects of semantic, morphological and syntactic violations. *Cognitive Brain Research*, *1*, 183–192.

Friedrich, M. & Friederici, A. D. (2004). N400-like semantic incongruity effect in 19-month-olds: Processing known words in picture contexts. *Journal of Cognitive Neuroscience*, *16*, 1465–1477.

Friedrich, M. & Friederici, A. D. (2005a). Phonotactic knowledge and lexical-semantic priming in one-year-olds: Brain responses to words and nonsense words in picture contexts. *Journal of Cognitive Neuroscience*, *17*(11), 1785–1802.

Friedrich, M.,& Friederici, A. D. (2005b). Lexical priming and semantic integration reflected in the ERP of 14-month-olds. *NeuroReport*, *16*(6), 653–656.

Friedrich, M. & Friederici, A. D. (2005c). Semantic sentence processing reflected in the event-related potentials of one- and two-year-old children. *NeuroReport*, *16*(6), 1801–1804.

Friedrich, M. & Friederici, A. D. (2006). Early N400 development and later language acquisition. *Psychophysiology*, *43*, 1–12.

Friedrich, M., Weber, C., & Friederici, A. D. (2004). Electrophysiological evidence for delayed mismatch response in infants at-risk for specific language impairment. *Psychophysiology*, *41*, 772–782.

Gout, A., Christophe, A., & Morgan, J. L. (2004). Phonological phrase boundaries constrain lexical access II. Infant data. *Journal of Memory and Language*, *51*(4), 548–567.

Gratton, G., Coles, M. G. H., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, *55*, 468–484.

Grimm, H. (2000). *Sprachentwicklungstest für zweijährige Kinder. Diagnose rezeptiver und produktiver Sprachverarbeitungsfähigkeiten*. Göttingen: Hogrefe.

Grimm, H. & Doil, H. (2000). *Elternfragebögen für die Früherkennung von Risikokindern (ELFRA-1, ELFRA-2)*. Göttingen: Hogrefe.

Guttorm, T. K., Leppänen, P. H. T., Poikkeus, A-M., Eklund, K. M., Lyytinen, P., & Lyytinen, H. (2005). Brain event-related potentials measured at birth predict later language development in children with and without familial risk for dyslexia. *Cortex. Special Issue: The neurobiology of developmental disorders*, *41*(3), 291–303.

Hahne, A. & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis. Early automatic and late controlled processes. *Journal of Cognitive Neuroscience, 11,* 194–205.

Hahne, A. & Friederici, A. D. (2002). Differential task effects on semantic and syntactic processes as revealed by ERPs. *Cognitive Brain Research*, *13*, 339–356.

Hahne, A., Eckstein, K., & Friederici, A. D. (2004). Brain signatures of syntactic and semantic processes during children's language development. *Journal of Cognitive Neuroscience*, *16*, 1302–1318.

Hebden, J. C. (2003). Advances in optical imaging of the newborn infant brain. *Psychophysiology*, *40*, 501.

Hillyard, S. A., Vogel, E. K., & Luck, S. J. (1998). Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *353*(1373), 1257–1270.

Hirsh-Pasek, K., Kemler-Nelson, D. G., Jusczyk, P. W., Cassidy, K. W., Druss, B., & Kennedy, L. (1987). Clauses are perceptual units for young infants. *Cognition, 26*(3), 269–286.

Holcomb, P. J. (1993). Semantic priming and stimulus degradation: Implications for the role of the N400 in language processing. *Psychophysiology*, *30*, 47–61.

Holcomb, P. J. & Neville, H. J. (1990). Auditory and visual semantic priming in lexical decision: A comparison using event-related brain potentials. *Language and Cognitive Processes*, *5,* 281–312.

Holcomb, P. J., Coffey, S. A., & Neville, H. J. (1992). Visual and auditory sentence processing: A developmental analysis using event–related brain potentials. *Developmental Neuropsychology*, *8,* 203–241.

Holm, A., Ranta-Aho, P. O., Sallinen, M., Karjalainen, P. A., & Müller, K. (2006). Relationship of P300 single-trial responses with reaction time and preceding stimulus sequence. *International Journal of Psychophysiology*, *61,* 244–252.

Houston, D. M., Santelmann, L. M., & Jusczyk, P. W. (2004). English-learning infants' segmentation of trisyllabic words from fluent speech. *Language and Cognitive Processes, 19*(1), 97–136.

Iacono, W. G., & Lykken, D. T. (1981). Two-year retest stability of eye tracking performance and a comparison of electro-oculographic and infrared recording techniques: Evidence of EEG in the electro-oculogram. *Psychophysiology*, *18*(1), 49–55.

Ille, N., Berg, P., & Scherg, M. (1997). A spatial components method for continuous artifact correction in EEG and MEG. *Biomedical Techniques and Biomedical Engineering*, *42* (Suppl. 1), 80–83.

Jasper, H. H. (1958). The ten-twenty electrode system of the international federation. *Electroencephalography and Clinical Neurophysiology*, *10,* 371–375.

Jing, H. & Benasich, A. A. (2006). Brain responses to tonal changes in the first two years of life. *Brain & Development, 28*(4), 247–256.

Jung, T.-P., Makeig, S., Westerfield, W., Townsend, J., Courchesne, E., & Sejnowski, T. J. (2000). Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. *Clinical Neurophysiology*, *111*(10), 1745–1758.

Jung, T.-P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., & Sejnowski, T. J. (2001). Analysis and visualization of single-trial event-related potentials. *Human Brain Mapping*, *14*, 166–185.

Jusczyk, P. W. (1997). *The discovery of spoken language*. Cambridge MA: The MIT Press.

Jusczyk, P. W., Cutler, A., & Redanz, N .J. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development*, *64*, 675–687.

Jusczyk, P. W., Hirsh-Pasek, K., Nelson, D. G., Kennedy, L. J., Woodward, A., & Piwoz, J. (1992). Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology, 24,* 252–293.

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, *39*(3–4), 159–207.

Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, *15*, 159–201.

Katayama, J. & Polich, J. (1998). Stimulus context determines P3a and P3b. *Psychophysiology*, *35*(1), 23–33.

Kenemans, J. L., Molenaar, P. C. M., Verbaten, M. N., & Slangen, J. L. (1991). Removal of the ocular artifact from the EEG: A comparison of time and frequency domain methods with simulated and real data. *Psychophysiology, 28*(1), 114–121.

Klann-Delius, G. (1999). *Spracherwerb*. Stuttgart: Metzler.

Kooijman, V., Hagoort, P., & Cutler, A. (2005). Electrophysiological evidence for prelinguistic infants' word recognition in continuous speech. *Cognitive Brain Research, 24*, 109–116.

Kooijman, V., Hagoort, P., & Cutler, A. (2006, March). Word recognition in continuous speech by 7-month-old infants. Paper presentation at the CUNY workshop On-line Methods in Children's Language Processing, New York, NY.

Kushnerenko, E., Cheour, M., Ceponiene, R., Fellman, V., Renlund, M., Soininen, K., et al. (2001). Central auditory processing of durational changes in complex speech patterns by newborns: An event-related brain potential study. *Developmental Neuropsycholgy, 19*(1), 83–97.

Kushnerenko, E., Ceponiene, R., Balan, P., Fellman, V., Huotilainen, M., & Näätänen, R. (2002). Maturation of the auditory event-related potentials during the first year of life. *NeuroReport*, *13*, 47–51.

Kutas, M. & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Science*, *4*, 463–470.

Kutas, M. & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207,* 203–205.

Kutas, M. & Hillyard, S. A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. *Memory & Cognition*, *11*(5), 539–550.

Leonard, L. B. (1998). *Children with specific language impairment*. Cambridge MA: The MIT Press.

Leppänen, P. H. T., Pikho, E., Eklund, K. M., & Lyytinen, H. (1999). Cortical responses of infants with and without a genetic risk for dyslexia: II. Group effects. *NeuroReport*, *10*, 969–973.

Leppänen, P. H. T., Richardson, U., Pikho, E., Eklund, K. M., Guttorm, T. K., Aro, M., & Lyytinen, H. (2002). Brain responses to changes in speech sound durations differ between infants with and without familial risk for dyslexia. *Developmental Neuropsychology, 22*, 407–422.

Lins, O. G., Picton, T. W., Berg, P., & Scherg, M. (1993). Ocular artifacts in EEG and event-related potential. II. Source dipoles and source components. *Brain Topography*, *6*, 65–78.

Lopes da Silva, F. (1991). Neural mechanisms underlying brain waves: From neural membranes to networks. *Electroencephalography and Clinical Neurophysiology*, *79*, 81–93.

Mangun, G. R. (1995). Neural mechanisms of visual selective attention. *Psychophysiology*, *32*, 4–18.

Marcar, V. L., Strassle, A. E., Loenneker, T., Schwarz, U., & Martin, E. (2004). The influence of cortical maturation on the BOLD response: An fMRI study of visual cortex in children. *Pediatric Research*, *56*(6), 967–974.

Martin, E., Joeri, P., Loenneker, T., Ekatodramis, D., Vitacco, D., Henning, J., & Marcar, V. L. (1999). Visual processing in infants and children studied using functional MR. *Pediatric Research*, *46*(2), 135–140.

Meek, J. (2002). Basic principles of optical imaging and application to the study of infant development. *Developmental Science*, *5*(3), 371–380.

Mills, D. L., Coffey-Corina, S. A., & Neville, H. J. (1993). Language acquisition and cerebral specialization in 20-month-old infants. *Journal of Cognitive Neuroscience*, *5*, 317–334.

Mills, D. L., Coffey-Corina, S. A., & Neville, H. J. (1994). Variability in cerebral organization during primary language acquisition. In G. Dawson & K. W. Fischer (Eds.), *Human behavior and the developing brain* (pp. 427–455). New York NY: Guilford Press.

Mills, D. L., Prat, C., Zangl, R., Stager, C. L., Neville, H. J., & Werker, J. F. (2004). Language experience and the organization of brain activity to phonetically similar words: ERP evidence from 14- and 20-month-olds. *Journal of Cognitive Neuroscience*, *16*, 1452–1464.

Morr, M. L., Shafer, V. L., Kreuzer, J., & Kurtzberg, D. (2002). Maturation of mismatch negativity in infants and pre-school children. *Ear and Hearing, 23*, 118–136.

Mrzljak, L., Uylings, H. B. M., Van Eden, C. G., & Judas, M. (1990). Neuronal development in human prefrontal cortex in prenatal and postnatal stages. *Progress in Brain Research, 85*, 185–222.

Näätänen, R. (1990). The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral and Brain Sciences*, *13,* 201–288.

Nazzi, T., Dilley, L. C., Jusczyk, A. M., Shattuck-Hufnagel, S., & Jusczyk, P. W. (2005). English-learning infants' segmentation of verbs from fluent speech. *Language and Speech*, *48,* 279–298.

Nazzi, T., Iakimova, G., Bertoncini, J., Frédonie, S., & Alcantara, C. (2006). Early segmentation of fluent speech by infants acquiring French: Emerging evidence for crosslinguistic differences. *Journal of Memory and Language*, *54,* 283–299.

Nelson, C. A. & Luciana, M. (1998). Electrophysiological studies II: Evoked potentials and event-related potentials. In C. E. Coffey & R. A. Brumback (Eds.), *Textbook of pediatric neuropsychiatry* (pp. 331–356). Washington DC: American Psychiatric Press.

Nobre, A. C. & McCarthy, G. (1994). Language-related ERPs: Scalp distributions and modulation by word type and semantic priming. *Journal of Cognitive Neuroscience, 6*(33), 233–255.

Oberecker, R. & Friederici, A. D. (2006). Syntactic event-related potential components in 24-month-olds' sentence comprehension. *NeuroReport*, *17*(10), 1017–1021.

Oberecker, R., Friedrich, M., & Friederici, A. D. (2005). Neural correlates of syntactic processing in two-year-olds. *Journal of Cognitive Neuroscience*, *17*, 407–421.

Opitz, B., Mecklinger, A., Cramon, D. Y., & Kruggel, F. (1999). Combining electrophysiological and hemodynamic measures of the auditory oddball. *Psychophysiology*, *36,* 142–147.

Osterhout, L. & Holcomb, P. J. (1992). Event-related potentials and syntactic anomaly. *Journal of Memory and Language*, *31*, 785–804.

Osterhout, L. & Holcomb, P. J. (1993). Event-related brain potentials and syntactic anomaly: Evidence on anomaly detection during perception of continuous speech. *Language and Cognitive Processes*, *8*, 413–437.

Pannekamp, A., Toepel, U., Alter, K., Hahne, A., & Friederici, A. D. (2005). Prosody-driven sentence processing: An event-related brain potential study. *Journal of Cognitive Neuroscience*, *17*, 407–421.

Pannekamp, A., Weber, C., & Friederici, A. D. (2006). Prosodic processing at the sentence level in infants. *NeuroReport*, *17*, 675–678.

Pascual-Marqui, R. D. (2002). Standardized low resolution electromagnetic tomography (sLORETA): Technical details. *Methods & Findings in Experimental & Clinical Pharmacology, 24*, 5–12.

Pascual-Marqui, R. D., Michel, C. M., & Lehmann, D. (1994). Low resolution electromagnetic tomography: A new method for localizing electrical activity in the brain. *International Journal of Psychophysiology, 18*, 49–65.

Pihko, E., Leppänen, P. H. T., Eklund, K. M., Cheour, M., Guttorm, T. K., & Lyytinen, H. (1999). Cortical responses of infants with and without a genetic risk for dyslexia: I. Age effects. *NeuroReport*, *10*, 901–905.

Ponton, C. W., Eggermont, J. J., Khosla, D., Kwong, B., & Don, M. (2002). Maturation of human central auditory system activity: separating auditory evoked potentials by dipole source modeling. *Clinical Neurophysiology*, *113*, 407–420.

Pujol, J., Soriano-Mas, C., Ortiz, H., Sebastián-Gallés, N., Losilla, J. M., & Dues, J. (2006). Myelination of language-related areas in the developing brain. *Neurology, 66*, 339–343.

Redcay, E., Haist, F., & Courchesne, E. (2006, March). Speech perception recruits frontal cortex during a pivotal period of language acquisition. Paper presentation at the CUNY workshop On-line Methods in Children's Language Processing, New York NY.

Regan, D. (1989). *Human brain electrophysiology: Evoked potentials and evoked magnetic fields in science and medicine*. New York NY: Elsevier.

Rescorla, L., Bascome, A., Lampard, J., & Feeny, N. (2001). Conversational patterns in late talkers at age 3. *Applied Psycholinguistics, 22,* 235–251.

Rescorla, L., Roberts, J., & Dahlsgaard, K. (1997). Late talkers at 2 outcome at age 3. *Journal of Speech, Language and Hearing Research*, *40*, 556–566.

Rivkin, M. J., Wolraich, D., Als, H., McAnulty, G., Butler, S., Conneman, N., et al. (2004). Prolonged T*[2] values in newborn versus adult brain: Implications for fMRI studies of newborns. *Magnetic Resonance in Medicine*, *51*(6), 1287–1291.

Rossi, S., Gugler, M. F., Hahne, A., & Friederici, A. D. (2005). When word category information encounters morphosyntax: An ERP study. *Neuroscience Letters*, *384,* 228–233.

Schapiro, M. B., Schmithorst, V. J., Wilke, M., Byars Weber, A., Strawsburg, R. H., & Holland, S. K. (2004). BOLD fMRI signal increases with age in selected brain regions in children. *NeuroReport*, *15*(17), 2575–2578.

Scherg, M. & von Cramon, D. (1986). Evoked dipole source potentials of the human auditory cortex. *Electroencephalography and Clinical Neurophysiology*, *65*, 344–360.

Scherg, M., Vajsar, J., & Picton, T. W. (1989). A source analysis of the human auditory evoked potentials. *Journal of Cognitive Neuroscience*, *1*, 336–355.

Silva-Pereyra, J., Klarman, L., Lin, L. J., & Kuhl, P. K. (2005a). Sentence processing in 30-month-old children: An event-related potential study. *NeuroReport*, *16*, 645–648.

Silva-Pereyra, J., Rivera-Gaxiola, M., & Kuhl, P. K. (2005b). An event-related brain potential study of sentence comprehension in preschoolers: Semantic and morphosyntactic processing. *Cognitive Brain Research, 23,* 247–258.

Soderstrom, M., Nelson, D. G. K., & Jusczyk, P. W. (2005). Six-month-olds recognize clauses embedded in different passages of fluent speech. *Infant Behavior & Development, 28*(1), 87–94.

Soderstrom, M., Seidl, A., Nelson, D. G. K., & Jusczyk, P. W. (2003). The prosodic bootstrapping of phrases: Evidence from prelinguistic infants. *Journal of Memory and Language*, *49*(2), 249–267.

Speckmann, E.-J. & Elger, C. E. (1993). Neurophysiological basis of the EEG and of DC potentials. In E. Niedermeyer & F. Lopes Da Silva (Eds.), *Electroencephalography. Basic principles, clinical applications and related fields* (pp. 15–26). Baltimore MD: Urban & Schwarzenberg.

Steinhauer, K., Alter, K., & Friederici, A. D. (1999). Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neuroscience, 2*, 191–196.

Szagun, G. (2006). *Sprachentwicklung beim Kind*. Weinheim: Beltz.

Tallal, P., Miller, S. L., Bedi, G., Wang, X., Nagarajan, S. S., Schreiner, et al. (1996). Language comprehension in language-learning-impaired children improved with acoustically modified speech. *Science, 271*(5245), 81–84.

Thierry, G., Vihman, M., & Roberts, M. (2003). Familiar words capture the attention of 11-month-olds in less than 250 ms. *NeuroReport*, *14,* 2307–2310.

Trainor, L., McFadden, M., Hodgson, L., Darragh, L., Barlow, J., Matsos, L., & Sonnadara, R. (2003). Changes in auditory cortex and the development of mismatch negativity between 2 and 6 months of age. *International Journal of Psychophysiology*, *51*, 5–15.

Uylings, H. M. B. (2006). Development of the human cortex and the concept of 'critical' or 'sensitive' periods. *Language Learning, 56* (s1), 59–60.

Weber, C., Hahne, A., Friedrich, M., & Friederici, A. D. (2004). Discrimination of word stress in early infant perception: Electrophysiological evidence. *Cognitive Brain Research*, *18*, 149–161.

Weber, C., Hahne, A., Friedrich, M., & Friederici, A. D. (2005). Reduced stress pattern discrimination in 5-month-olds as a marker of risk for later language impairment: Neurophysiologial evidence. *Cognitive Brain Research, 25*, 180–187.

West, W. C. & Holcomb, P. J. (2002). Event-related potentials during discourse-level semantic integration of complex pictures. *Cognitive Brain Research*, *13*, 363–375.

Winkler, I., Kujala, T., Tiitinen, H., Sivonen, P., Alku, P., Lehtokoski, A., et al. (1999). Brain responses reveal the learning of foreign language phonemes. *Psychophysiology*, *36*(5), 638–642.

Woldorff, M. G., Gallen, C. C., Hampson, S. R., Hillyard, S. A., Pantev, C., Sobel, D., & Bloom, F. E. (1993). Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proceedings of the National Academy of Sciences*, *90*, 8722–8726.

Appendix to Chapter 2

1. How to run an ERP experiment with infants

In working with infants, we are faced with certain limitations that make the experimental procedure much more challenging than with adults. Specifically, an abbreviated attention span, limited verbal and motor skills, frequently occurring hunger and tiredness necessitate short experiments that work without instructions and do not require verbal or motor responses. Ideally, passive listening paradigms (like the *passive oddball paradigm* and *semantic/syntactic violation paradigm*), sometimes combined with passive viewing (like the *picture-word paradigm*), are used and adjusted to the infants' abilities. In our laboratory, we find that experiments should take no longer than 10–15 minutes; unless the outcome is not affected by the infants being asleep (in this case we have found 20–30 minutes acceptable). Although the trial number has to be drastically reduced for this reason, the presentation of at least 30–60 trials per condition is required for subsequent averaging. This can be done by reducing the number of experimental conditions and simplifying the stimulus material as compared to adult experiments. To ensure a high quality of the recorded EEG data, we also find it useful to program the experiment so that it can be paused at any point in time or after certain experimental blocks. This allows the experimenter to immediately react to the infants' needs, such as rearranging their seating or reaching for a pacifier.

Passive listening and viewing paradigms enable comparisons across age groups because infants, children, and adults can all be tested under the same conditions. For adults, the passive testing condition is sometimes difficult because it requires maintaining attention throughout the experiment without performing a task. Therefore, in adults it might be useful to monitor the subjects' attention via video and motivate them through general instructions and the announcement of a post-experimental questionnaire on the experiment. Accordingly, possible attention-related differences in testing conditions should be taken into consideration when comparing data across age groups or data of different studies.

2. Hands on infants' brains: EEG recordings

During the set-up procedure, the infants usually sit on their parents' lap, watching a book or playing. While preparing the infants for the EEG recording, all steps should be simultaneously explained to the parents. The set-up time is individually variable (e.g., dependent on how shy infants are or whether they are familiar with the situation) and is also dependent on the infants' age. The younger the infants (1–6 months), the faster the set-up should be completed to avoid early fatigue effects during the experiment. From about 4 months on it becomes more and more important to offer interesting distractions, such as jingling toys, hand puppets, or movies. From about 18 months on, the experimental procedure should be additionally explained to the children – convincing them on how 'fancy' and 'cool' it is to wear such a hat or letting them do the same set-up on a doll. Since this work demands attention to both infants and parents and requires very flexible arrangements, we recommend that two experienced experimenters are present. Also, the set-up should take place in a preparation room (other than the testing room) with lots of toys and baby equipment to make infants and parents feel comfortable.

Our EEG recording equipment consists of differently sized elastic caps fitted with a low-density montage of silver-silverchlorid electrodes (*EASY CAP GmbH*). We find that this low-density montage system fits the requirements of our developmental research because it is easy to set up and we do not aim to perform source localization (for source localization purposes see high-density EEG recordings by *Geodesic Sensor Net, Inc.*). The number of electrodes used varies from 9 to 32 depending mainly on the infants' age (head size) and the applied paradigm. All electrodes are arranged according to the International 10–20 System of Electrode Placement (Jasper 1958) or the extended 10–10 System (Chatrian, Lettich, & Nelson 1988), and are attached to the cap by ring adapters. This montage technique allows for simple cleaning by rinsing cap and electrodes separately in warm water and for easy exchange of broken electrodes. To monitor eye movements (see data analysis), we place additional electrodes on the outer canti of both eyes and on the infra- and supraorbital ridges of the right eye (electrooculogram; EOG). Note that, except the electrode below the right eye, all electrodes (including common reference and ground) are implemented in the cap, so that only one single electrode has to be placed at the infant's head, which allows an easy and quick set-up. To ensure proper electrode placement the elastic cap is fixed by a chest strap because chin straps (often used in adults) are not comfortable for infants. To apply the conductance gel under the ring electrodes, we simply use a blunt tip syringe and cotton swabs. Note that the conductance gel should be non-abrasive and not contain any preservatives.

The EEG recording takes place in a noise-shielded cabin, only equipped with loudspeakers and a video screen to keep the infants focused on the experiment. During the recordings, the infants sit on their parents' lap or in a car seat with the parents sitting next to them (preferably the latter, since it minimizes movement). In either case, it is very important to carefully instruct the parents: no talking, no bouncing or swinging, and if necessary redirecting the infants' attention to the screen by pointing. Also, the parents should wear ear plugs or headphones so that they cannot hear the acoustic stimuli. During the whole time, the infants are monitored via video and if necessary the recording is paused. To achieve eye fixation and reduce movement we usually provide a silent video (e.g., slowly moving fish) and try to attract the infants' attention by hand puppets, soap bubbles, etc. when they become inattentive and restless. If necessary, the infants also may have a pacifier, bottle, etc. Given the variability in the infants' state, we highly recommend to carefully document the exact course of the experiment and the specific testing conditions to determine possible external influences, such as time of the day, nursing, tiredness, crying, use of bottle or pacifier, kind of distraction.

3. What to do with the recorded EEG: EEG data analysis

The off-line processing of the recorded EEG data focuses on the time-locked averaging of the events of interest to derive ERPs. In some cases, however, specifically in 'noisy' infant data, prior filtering and artifact reduction procedures are unavoidable to remove unwanted noise, caused by head and eye movement, perspiration, etc.

Dependent on the electrode referencing during the recording, the data may have to be re-referenced. In our experiments, the EEG is recorded to the common reference Cz (an electrode that is about equidistant from all other electrodes and enables stable recording). Off-line, the data are then re-referenced to a mathematically linked reference from two separately recorded electrode sites that are somewhat distant from the other sites and relatively 'inactive', e.g., the mastoids. It is important to note that the choice of reference has an influence on the amplitudes as well as the topography of the ERPs and should be considered when comparing data of different studies (for discussion see Dien 1998).

Filtering refers to the removal of certain frequencies from the signal that are sufficiently different from the frequencies that contribute to the ERP wave form (a matter that becomes more complicated as the frequency content of the noise and the signal are more similar). Accordingly, most researchers apply band pass filters that spare, for instance, frequencies between

0.01 Hz and 30 Hz, known to reflect most of the portion of the ERP wave form. In this way, high frequencies (e.g., from muscle contractions) as well as very slow frequencies (e.g., from skin potentials/perspiration) are suppressed. The latter is especially relevant for infant data that are often drift-contaminated and should not enter ERP averages without prior filtering. Here, high-pass filters might be even set to 0.5 Hz or 1.0 Hz, if the experiment does not focus on the analysis of slow potentials. Thus, filters should be designed having both the features of the recorded signal and the expected ERP results in mind. In other words, we generally recommend the use of filters but also advise their cautious application because filters can substantially distort the ERP data (for more detail on the function and design of filters see, for example, Edgar, Stewart, & Miller 2005).

Artifact rejection procedures are applied to increase the signal-to-noise ratio by eliminating trials that contain artifacts. Most artifacts are substantially larger in amplitude than the EEG signal and originate from various external sources, such as eye movement, head or body movement, and technical equipment. Automatic artifact rejection procedures are for instance set to exclude all trials exceeding 100 μV (in adults) or 200 μV (in infants). However, given the high variability of infant data consideration, we recommend rejecting trials that exceed for instance a standard deviation of 80 μV in a sliding time window of 200 ms, rather than setting an absolute amplitude criterion.

A great deal of signal contamination stems from ocular artifacts, particularly blinks that primarily contribute to the EEG measured at anterior sites. Since infants, in contrast to adults, cannot be given 'blink instructions' the crude rejection of contaminated data might leave too few trials. For those cases, artifact correction procedures might be more appropriate than the complete exclusion of contaminated trials. There are various correction procedures that, in a first step, mathematically calculate the propagation between the eyes and each of the scalp electrodes or determine the spatial components of the eye activity. In the second step, the estimated EOG proportion (propagation factors) is subtracted from the EEG signal at each site (for a comparison of the different methods see Brunia et al. 1989). The correction techniques differ in their mathematical approach and the involved theoretical assumptions, namely regression methods (e.g., Gratton, Coles., & Donchin 1983; Kenemans, Molenaar, Verbaten, & Slangen 1991), dipole source modeling (e.g., Berg & Scherg 1994; Lins, Picton, Berg, & Scherg 1993), principal component analysis (PCA, e.g., Ille, Berg, & Scherg 1997), and independent component analysis (ICA, e.g., Jung et al. 2000). In the analysis of infant data, we highly recommend artifact correction procedures to avoid losing too many trials. However, it is important to keep in mind that each method has its own specific flaws, such as 'over-correction' of EEG signal in regression methods (since the EEG signal also propagates to EOG channels and contributes to the estimation of regression coefficients; see Iacono & Lykken 1981).

For the averaging procedure a minimum number of artifact-free trials is required. In our laboratory we require each infant to deliver at least 10–20 artifact-free trials to the individual average and at least 20–30 subjects to contribute to the overall average. Thus, to complete an experiment we usually test 30–50 infants with a success rate of 65–85%, depending on the their age (e.g., 4-month-olds are easier to test than 12-month-olds, who want to get up and walk around) and the demands of the experiment (e.g., passive listening alone often has a higher success rate than additional passive viewing).

CHAPTER 3

Using eye movements as a developmental measure within psycholinguistics

John C. Trueswell

This chapter describes and evaluates the use of eyetracking methods to study the development of spoken language production and comprehension. The emphasis will be on understanding the chain of inferences, or linking assumptions, researchers commonly make when going from measurements of eye position to conclusions about attention, reference and sentence parsing. It is argued that these assumptions are valid, though care is needed when disentangling developmental changes in visual attention from developmental changes in language processing abilities.

1. Introduction

Cognitive development is often viewed as the acquisition of knowledge: we learn facts about the world around us; we learn the words and the grammar of a language, etc. An alternative way of thinking about cognitive development, which has gained some traction recently in the developmental literature, is to treat it as the acquisition of dynamic skills: we learn how to interact with the world; we learn how to produce and comprehend a language, etc. The work discussed in this volume is about this dynamic processing approach to development, particularly as it pertains to language development. Recent interest in this issue stems in part from concurrent methodological advancements; it is now possible for instance to record children's eye movements as they carry out relatively natural tasks involving language, such as following spoken instructions, inspecting images that are being described, and even engaging in a spoken conversation with interlocutors. The resulting eye movements, when linked with linguistic events, provide researchers with a record of each child's moment-by-moment consideration of possible referents in the world and thus tell us in some detail about the process the child is going through when deriving meaning from linguistic forms.

This chapter describes and evaluates this “visual world” method (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy 1995) and focuses especially on how it has been applied to sentence processing research with toddlers and children. The emphasis here will be on understanding the linking assumptions necessary to use eye movements to study language development. That is, this chapter will explore the chain of inferences researchers usually make when going from measurements of darting eyes to conclusions about attention, reference and even sentence parsing. The plan is to step through these linking assumptions and explore the extent to each is valid and how each might interact with known developmental changes in attention.

I hope to convince the reader that conclusions drawn from developmental research using this visual world paradigm require careful consideration of how certain attentional skills develop, in particular, the developing abilities to engage in the control of information collection (a component of *attentional control*) and information re-characterization (a component of *cognitive control*). I will discuss how these two kinds of attentional abilities change over development, and how these changes might bear upon the interpretation of eye movement research in psycholinguistics. With respect to information collection, it is well known that the eye movements generated during the visual interrogation of the world are driven by both exogenous and endogenous factors (i.e., by both bottom-up visual factors and experience-related goals set by the individual). With respect to information re-characterization, it is well known that humans routinely characterize perceptual input along several different dimensions at several levels of abstraction. Language is perhaps the parade example of this; we characterize linguistic input acoustically, phonologically, syntactically, semantically and referentially, with each characterization having its own representational dimensions. Adult listeners must be able to control the content of these characterizations in real-time and override certain characterizations when conflicting evidence arises within and across these levels. Indeed, the skill of dealing with conflict turns out to be important in the development of sentence comprehension abilities.

With this broader understanding of how attentional and cognitive control abilities develop, researchers are likely to make (and are already making) significant advances in understanding how the dynamics of language comprehension and production emerge in the young child. It is my hope that touring these facts here will allow others to take advantage of the visual world method, and that it will facilitate theoretical advancements in understanding language acquisition as the development of a dynamic information processing skill.

2. The visual world paradigm and developmental psycholinguistics

Cooper (1974) was the first to use eye movements as a real-time measure of adult's spoken language processing abilities. In a series of eyetracking experiments, it was observed that adult listeners rapidly fixate pictures depicting the referents of heard speech, often mid-word, prior to the completion of the utterance. This work received fairly limited discussion in the psycholinguistic community until the re-introduction of this method by Tanenhaus and colleagues who explored the eye gaze of listeners in the natural setting of following spoken instructions to move about objects in the world (Tanenhaus et al. 1995). Tanenhaus et al. demonstrated that when adult participants follow spoken instructions to manipulate objects in a task-relevant visual context, fixations to these objects are also closely time-locked to the elements present in the unfolding utterance that signal abstract representational units. It was therefore possible from this work to infer a great deal about the lexical (e.g., Allopenna, Magnuson, & Tanenhaus 1998) and syntactic (e.g., Spivey, Tanenhaus, Eberhard, & Sedivy 2002) hypotheses that adults consider as the speech is perceived. Since publication of this seminal work, a growing body of research has demonstrated that eye movements can be used to trace the time course of adult language comprehension, production and even dynamic conversation. (See the edited volumes of Henderson & Ferreira 2004; and Trueswell & Tanenhaus 2005; for thorough reviews.)

2.1 Eyetracking techniques for use with children

The development of accurate *head-mounted* and *remote* eyetracking systems has made it possible to conduct similar visual world studies with young children, toddlers and even infants. Head-mounted systems (Figure 1) use highly miniaturized cameras and optics mounted on a visor (two cameras, one trained on the eye and the other on the surrounding visual world). In these systems, the video output from the eye camera is analyzed in real time to calculate the current location of the pupil (i.e., the central position of all the darkest pixels) and the center of the corneal reflection (i.e., the central position of the brightest pixels). During an initial calibration procedure, these coordinates are mapped onto coordinates in the scene video. This is typically done by asking the participant to look at locations in the world that correspond to particular pixel coordinates in the scene video. For each location, the pupil and corneal reflection coordinates in the eye camera are sampled and paired with a coordinate position in the scene camera. (Informally, the computer is being told that the participant's eyeball looks like *this* when the participant is looking *here* and it looks like *this* when the participant is looking over *here*, etc.). The

Figure 1. Head-mounted eyetracking.

resulting matrix of coordinates (triplets of pupil, corneal reflection and position coordinates) is then analyzed. This analysis creates a multi-dimensional linear or nonlinear regression equation that reflects the best fit between the eye-calibration coordinates and the scene-calibration coordinates. This equation can then be applied in real time throughout the experiment, such that for any pupil and corneal coordinates, the corresponding scene coordinate is generated and plotted on top of the scene video (usually as a moving dot or crosshair).

This calibration procedure can be difficult to use with children because it requires the child to hold his/her head still while fixating a target location in the world. However, some calibration procedures eliminate this problem; in the point-of-light calibration procedure, the experimenter holds a small light (such as a small LED) while the participant follows the light around with his/her eyes. The eyetracking calibration software then samples the position of this bright light in the scene video and pairs it with the pupil and corneal coordinates from the eye video, thereby creating a calibration matrix. This procedure does not require the child to hold still, and substantially decreases calibration time and increases calibration accuracy.

Remote eyetracking systems (Figure 2) work like head-mounted systems except the optics are housed off the head, requiring no visor. These systems require tracking of the head as well, either via video-based methods (e.g., the Tobii 1750) or by magnetic head tracking (e.g., the ASL and ISCAN systems). Remote systems are becoming increasingly popular because they can be easier to use with toddlers and even infants (e.g., Aslin & McMurray 2004; S. Johnson, Slemmer, & Amso 2004). Most remote systems map direction of gaze directly onto the coordinates of a computer video display, rather than a scene camera, allowing for simple automatic coding of eye position. However, it is also possible to use such systems to generate a three-dimensional vector of the participant's gaze in the physical world rather than a virtual world.

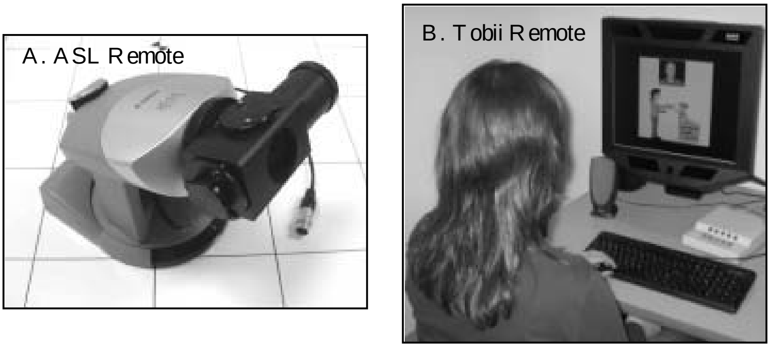


Figure 2. Remote eyetracking.

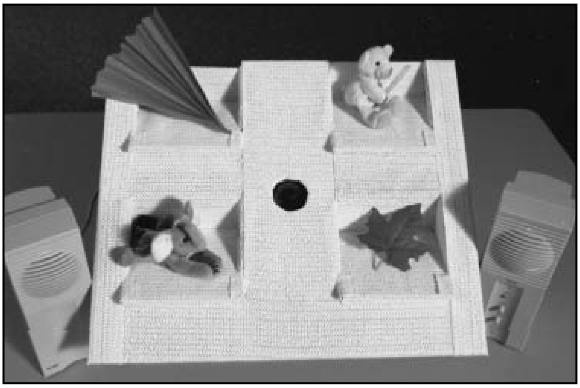


Figure 3. Poor man's eyetracker.

Finally, several labs (including my own) also use what we affectionately call the "poor man's eyetracker" (Figure 3, see also Snedeker & Thothathiri this volume). In a modified preferential looking procedure, a video camera is located in the center of a platform that has been placed in front of the child. This camera is trained on the child's face and eyes. Objects are placed on the platform, usually in four different quadrants around the camera. Direction of gaze toward each quadrant can be coded from the video of the child's face; a trained coder can use a digital video editing system to step through the video frame-by-frame, recording shifts in gaze. Hand coding of this sort is quite time consuming; it takes approximately an hour to code ten to fifteen experimental trials when each trial consists of one or two utterances. However, no calibration procedure or expensive eyetracking equipment is required. This hand-coding procedure also tolerates considerable head movements without substantial loss in coding accuracy. We have found that inter-coder reliability is usually 90–95% on a frame-by-frame basis (Snedeker & Trueswell 2004). Similar hand-coding procedures are used in picture-viewing tasks with infants and toddlers (see, e.g., Fernald, Zangl, Portillo, & Marchman this volume; Swingley, Pinto, & Fernald 1999).

2.2 Data analysis

Regardless of the data collection technique used by the experimenter, similar analyses can be performed on the resulting gaze record. For each trial of interest, the child's direction of gaze is linked to the onset of critical speech events (e.g., the onset of critical words in a sentence) and then averaged across trials and participants. For example, Trueswell, Sekerina, Hill and Logrip (1999) evaluated the time course with which 5-year-old children visually inspect a set of four possible referents, relative to critical word onsets in a sentence. The children were instructed to look at a centrally located 'smiley face' sticker and then to follow instructions to move some of the objects. For purposes of illustration consider a hypothetical trial in which participants heard: *Look at the smiley face. Now put the frog that's on the napkin into the box.*

A photograph of a sample scene for this item is presented in Figure 4.[1] Objects include the target (a frog on a napkin), the competitor (a frog on a plate), a correct goal (an empty box) and an incorrect goal (an empty napkin). The upper right panel of Figure 4 shows the eye gaze records from five hypothetical trials. The zero ms point (where the x and y axes meet) indicates the onset of the spoken word *put*. In addition, the onsets of the nouns are marked (*frog*, *napkin* and *box*). On trial one, the hypothetical participant initiated a look to the target about 400 ms after the onset of the word *frog* and then launched a look to the correct goal later in the sentence. On trial two, the fixation on the target begins a bit later. On trial three, the first fixation is on the competitor, followed by a fixation on the target and then the correct goal. On trial four, the fixation sequence is target, incorrect goal, and correct goal. Trial five shows another trial where the initial fixation is on the competitor. The lower right panel of Figure 4 provides a plot of the proportion of looks over time for the four regions, averaged across trials for this hypothetical participant. These fixation proportions are obtained by determining the proportion of looks to the alternative objects at each time slice (as derived from the trial samples) and show how the pattern of looks to objects changes as the sentence unfolds.[2] The probabilities do not sum to 1.0 because most participants were ini-

---

[1] This figure is modeled after a similar discussion of eye movements appearing in Tanenhaus and Trueswell (2005).

[2] Like most psycholinguistic studies, several similar target trials are provided to the same participant; e.g., in addition to the frog item, there might be an item involving cows: *Put the*

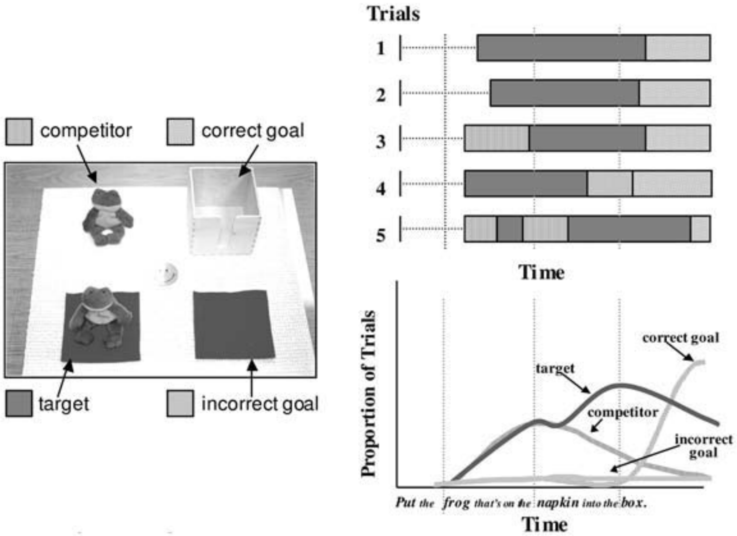


Figure 4. Calculating gaze proportions over time (modified from Tanenhaus & Trueswell 2005).

tially fixating on the smiley face, which is not plotted here. If it were plotted, looks to the smiley face would steadily drop over time while children begin to inspect the task-relevant objects.

Researchers often define a time window of interest. For example, one might want to focus on the looks to the target and competitor in a time region starting 200 ms after the onset of the word *frog* and ending 200 ms after the onset of the word *napkin*. This 200 ms offset is designed to take into account that it takes about 200 ms for a participant to program an eye movement to a target (e.g., Matin, Shao, & Boff 1993, though see below). The proportion of looks to objects, the time spent looking at the alternative objects (essentially the area under the curve, which is a simple transformation of proportion of looks), and the number and/or proportion of looks generated to objects in this time region can then be analyzed. These different measures are all highly correlated but in principle offer slightly different pictures of the eye movement record.

The actual data from this condition in the Trueswell et al. study are reproduced in Figure 5. Focusing only on the looks to the target and the competitor, one can see that these looks are fairly well time-locked with the onset of words;

---

cow that's in the box onto the book and so on. These target items are identical to each other in all experiment-relevant ways but differ in terms of the nouns and the objects they refer to. To avoid task demands, target trials are randomly mixed among numerous filler trials containing a variety of different linguistic structures and different referents.

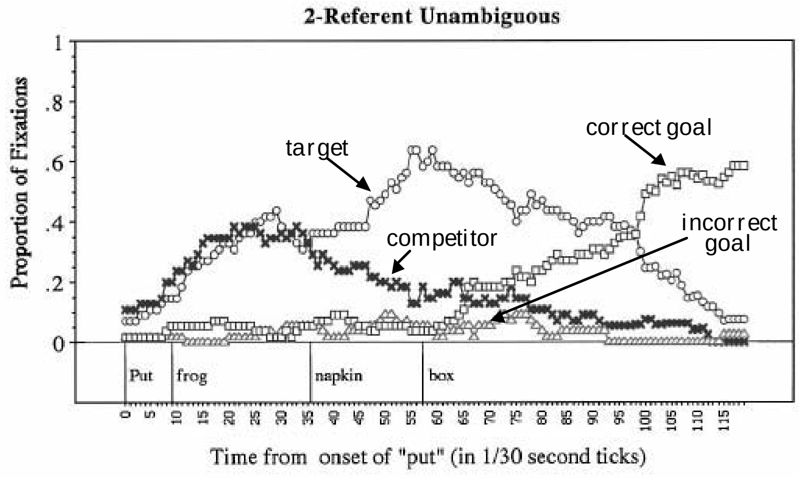


Figure 5. Children's (5 year olds') proportion of looks over time to potential referent objects in response to *Put the frog that's on the napkin into the box*. From Trueswell, Sekerina, Hill and Logrip (1999). Copyright 1999, Elsevier Press.

first, looks to both the target and competitor (the two frogs) rise sharply upon hearing the first noun, *frog*, and remain equally distributed between these two objects until *napkin*, at which time participants begin to look more at the target (the frog on the napkin). Similarly, looks to the correct goal rise upon hearing *box*.

It is not the case that the eyes simply dart to objects that best match the nouns mentioned in the input. For instance, at the onset of the noun *napkin*, eye position does not split between the two napkins in the scene like it did for the two frogs when hearing *frog*. Rather, looks to the target (that has the napkin under it) prevail over looks to the incorrect goal (the empty napkin). Why would this be? The most plausible explanation is that this is due to the syntactic position of the noun *napkin* in the sentence; this noun is part of a relative clause that unambiguously modifies the NP *the frog* (i.e., *the frog that's on the napkin*); as such, the NP *the napkin* must refer to the napkin under the frog, not the empty napkin. Similar time-course data has been reported for adults (e.g., Spivey et al. 2002; Tanenhaus et al. 1995; Trueswell et al. 1999) and replicated in other children (Hurewitz, Brown-Schmidt, Thorpe, Gleitman, & Trueswell 2001; Snedeker & Trueswell 2004), all of which suggests that gaze direction is tightly related to the linguistic events in complex sentences and that reference is being computed by the child and adult listener in real time.

3. Linking assumptions

It is crucial to consider the *linking assumptions*, or chain of inferences, that we just rapidly ran through when evaluating data like those in Figure 5. As compelling as this probability plot is, how can we confidently go from eye gaze patterns to the conclusion that child listeners compute referential hypotheses in real-time? In order to answer this question, there are at least three crucial linking assumptions worth evaluating further.

(1) Eye position indicates the child's current attentional state, and attention is driven by properties of the world and by the goals of the child.
(2) In tasks requiring the linking of speech to a visual referent world, visual attention can be used as an indication of referential decisions.
(3) Referential decisions can in turn be used by the researcher to infer the child's syntactic parsing decisions, insofar as these parsing decisions were necessary to determine the referent.

The remaining sections of this chapter unpack each of these linking assumptions and examine the current experimental literature for validation of these assumptions. In particular, Section 3.1 (*Eye position is a metric of spatial attention in adults, children and infants*) examines the first of these three linking assumptions. Section 3.2 (*Eye movements can be used to infer referential and syntactic decisions*) examines the latter two assumptions. Wherever relevant, these sections also explore how observed developmental patterns might interact with these linking assumptions and therefore modify the conclusions that can be drawn when using the visual world method in developmental psycholinguistics.

3.1 Eye position is a metric of spatial attention in adults, children and infants

3.1.1 *Ocular development*

In normal everyday visual inspection of the world, adults rapidly shift their eyes from location to location approximately 1 to 5 times per second. During these rapid eye movements, or saccades, the eye is in motion for 20 to 60 ms, and can reach speeds of 500 to 1000 degrees per second. Saccades allow for the repositioning of visual input onto the fovea, a small central region of the retina that, because of its higher density of photoreceptors, has considerably better image resolution than peripheral retinal regions. Each saccade is followed by a fixation, during which the eye holds essentially still for 150 ms or more depending on the task.

(For sources and references on eye movements, see Kowler 1995; Liversedge & Findlay 2001; Rayner 1998.)

For the normally developing newborn, most of these anatomical properties of the retina are in place at birth or develop rapidly during the first months of life. Basic fundamental oculomotor abilities are also in place quite early; saccades, fixations, and even the ability to smoothly pursue a slowly moving object all emerge quickly during the first six months of life and are known to be well in place by the child's first birthday (for a review, see Colombo 2001). As discussed below however, quantitative developmental changes in eye movement abilities do occur well after the first birthday.

3.1.2 *Saccade latency*

Only a small number of studies have examined how the latency to launch a saccade to a visual target changes with age in children (Cohen & L. Ross 1977, 1978; J. Fukushima, Hatta, & K. Fukushima 2000; S. Ross & L. Ross 1983; Salman et al. 2006; Yang, Bucci, & Kapoula 2002). All of these studies show that saccade latency steadily decreases well into the age ranges studied by most psycholinguists. For instance, Yang et al. (2002) report that saccade latency to a visually selected target is on average about 450 ms for 4.5-year-old children and decreases steadily with age to approximately 250 ms for 12-year-old children and adults. However, some of these developmental differences may be related to response preparation and/or the specifics of the task. Cohen and L. Ross (1978) report that 8-year-old children are as fast and accurate as adults in making saccades to a target when the target was preceded by a 300 ms warning (see also S. Ross & L. Ross 1983). This latter finding may be particularly relevant to the psycholinguistic visual world method because ample response 'warning' is given in this task, via linguistic input (*Look at the smiley face. Now put…*). Children (5-year-olds) in visual world tasks appear to show only modest delays in their latency to find a target (Snedeker & Trueswell 2004). And toddlers (18 months) show a 150 ms benefit in targeting a visual referent when the referential expression is preceded by a linguistic carrier phrase as compared to when it is not (Fernald & Hurtado 2006; see also Fernald, Zangl, Portillo, & Marchman this volume).

3.1.3 *Eye position as index of spatial attention*

Although adults can direct spatial attention to regions of space that are not currently being fixated (often called covert spatial attention, Posner 1980), a growing body of behavioral and neurophysiological work supports a close link between current fixation and spatial attention (Findlay 2004; Kowler 1995; Liversedge & Findlay 2001). Under this view, selection of an object for fixation is determined by a weighted combination of exogenous and endogenous factors. Attention is in part controlled exogenously, i.e., by visual properties of the world that 'capture' our attention. Regions of space that are highly distinguished from other areas (especially sudden onsets of motion, a.k.a., motion transients) draw our gaze quite rapidly and automatically (e.g., Franconeri & Simons 2003; Jonides & Yantis 1988; Yantis & Jonides 1984, 1990).[3] In contrast, experience-driven expectations and navigational plans also contribute to the visual selection of an object or a region of space, and thus contribute endogenously to attentional control. Under many natural viewing conditions, endogenous factors must override exogenous influences so as to allow for the guidance of attention to objects that are task-relevant but otherwise visually less salient (e.g., Guitton, Buchtel, & Douglas 1985; Hallett 1978).

There is still great debate in the attention literature regarding the details of how objects are selected for attention (e.g., in visual search tasks). However, many current neurocomputational models of attention propose a parallel selection process (e.g., Findlay 2004; Rao, Zelinsky, Hayhoe, & Ballard 2001; Zelinsky, Rao, Hayhoe, & Ballard 1997). Here, the entire visual field at any moment in time is characterized by multiple 'saliency' maps (color saliency, motion saliency, texture saliency, etc.). Goal-directed orienting is driven by active integration of these saliency maps; for instance, searching for a toy frog would be hypothesized to include parallel consultation of those spatial regions distinguished by relevant colors, motions, textures, etc. A viewer's memory for the spatial position of objects is also likely to play a role (e.g., a recent memory for the location of a toy frog). A viewer's working memory for the spatial position of objects appears to be limited to a small set of items (e.g., Pylyshyn 1994). However it is also believed that the extent to which a viewer holds the visual details of these objects in memory depends greatly on the task. For instance, in many simple manual tasks, the perceiver may rely on the external world as a kind of visual 'memory' by assuming that the visual features of objects in the world remain unchanged over time (e.g., Ballard, Hayhoe, & Pelz 1995).

Quite clearly, eye movement measurements from the visual world paradigm rely on the participant's accurate implementation of visual selection processes, for which task-relevant endogenous factors are expected to override exogenous factors: Participants are expected to fixate on what is relevant for carrying out the instruction, not on what is most colorful or eye-catching. In fact, in an important sense, the linguistic input can be viewed in this approach as a straightforward characterization of the participant's current spatial goals (*Look at the smiley face, Pick up the frog,* etc.). Given this, it seems urgent to understand the developmental time course of endogenous visual selection abilities, particularly in nonlinguistic

---

[3] See also Ruz and Lupiáñez (2002), Yantis (1993), and Yantis and Jonides (1996), for debates over what exactly captures attention and whether attention capture is truly automatic.

tasks. Without knowledge of these facts, one will not be able to adequately interpret developmental visual world findings, particularly as they apply to theories of language processing and language development.

As it turns out, endogenous control of attention by infants begins quite early in life but appears to have a protracted developmental profile. For instance, infants who are 3.5 to 6.0 months of age or older can learn to reposition their eyes so as to anticipate the appearance of an object in predictable locations (Haith, Hazan, & Goodman 1988; McMurray & Aslin 2004). These studies, however, do not involve procedures that require infants to override exogenous factors that might influence attention. To what extent can infants and toddlers do this? In the first study of this kind, M. Johnson (1995) placed 4-month-olds in a modified *anti-saccade* task: Children had to learn that the sudden onset of a spatial cue on the left predicted the onset of a rewarding visual stimulus on the right, and vice versa. Adults in such tasks (e.g., Hallett 1978; Guitton et al. 1985) rapidly learn simultaneously to inhibit attention shifts to the briefly presented spatial cue (i.e., to inhibit pro-saccades to the cue) and to generate anticipatory looks to the reward location (i.e., to generate anti-saccades, despite the sudden onset on the other side of the screen). M. Johnson (1995) found that 4-month-olds did learn over the course of the experiment to inhibit pro-saccades. That is, they learned to not look at the flashed spatial cue presumably because it became clear that this event was perfectly correlated with the presence of a reward on the other side of the screen. However, these same infants were never able to generate anti-saccades. That is, they did not learn to move their eyes to the reward location prior to its visual onset. Importantly, as we just mentioned, infants in this age range can anticipate the location of a reward object when a (non-anti-saccade) spatial cue is provided (i.e., the Haith et al. 1998, result). Taken together then, it appears that the best that a 4-month-old can do is counter-act but not completely override exogenous contributions to attention.

Recently, Scerif, Karmiloff-Smith, Campos, Elsabbagh, Driver and Cornish (2005) examined the development of anti-saccade abilities over a much larger age range (8 to 40 months). Scerif et al. found that the proportion of pro-saccades steadily decreases within this age range (from 100% down to approximately 20%) whereas the proportion of anti-saccades steadily increases (from 0% to approximately 40%). That is, it appears that the ability to simultaneously counter-act exogenous factors while promoting endogenous factors has a fairly protracted developmental profile; children under the age of three years are more susceptible to exogenous factors than older children.

3.1.4 *Implications for the psycholinguist*

What are the implications for those developmental psycholinguists who use eye movements to infer language processing abilities? First, it is clear that eye position is an excellent metric of attention, even in young children. Provided that children are given some warning that they will have to find a target, saccade latencies show only modest delays relative to adults. However, the data show that exogenous and endogenous contributions to attention change over the course of development, even into ages tested in many psycholinguistic studies (2.0–3.5 years). It is difficult to draw straightforward connections between the developmental attention literature and the developmental psycholinguistic literature because most psycholinguistic experiments use very different experimental settings. However, if the relative influence of exogenous and endogenous factors changes over developmental time, it becomes quite important for psycholinguistic researchers to control for visual factors known to capture attention (e.g., motion, sudden onsets). Otherwise, developmental changes that are simply due to general attentional development might be misinterpreted as developmental changes related to spoken language understanding.

For instance, many preferential looking studies compare how children's visual inspection of two side-by-side animated movies changes as a function of linguistic input and as a function of development, often between the ages of 1 to 4 years (e.g., Golinkoff, Hirsh-Pasek, Cauley, & Gordon 1987). This is precisely the age at which substantial developmental changes are occurring in attentional control. If differences in visual saliency within contrasting videos (especially differences in the timing of sudden-onsets and motion within these videos) are not carefully matched, it is possible to have a situation in which the youngest children's inability to use linguistic evidence to guide attention is actually due to an inability to override exogenous factors. An infant/toddler might understand the meaning of an utterance but fail to use this information to shift gaze because attention has been captured elsewhere by lower level visual properties. The use of control conditions, in which infant/toddler viewing patterns are recorded in the absence of relevant linguistic input, does help circumvent this problem, but only partially. It also needs to be established that for each pair of videos tested, no strong viewing preferences exist in the control condition.

3.2 Eye movements can be used to infer referential and syntactic decisions

We now turn to linking assumptions 2 and 3, which can be combined and restated more precisely as follows.

If a task requires linking speech to a visual referent world, eye movement experiments can be designed to uncover the listener's ongoing referential decisions and, by inference, their ongoing syntactic parsing decisions. Note that this does *not* mean that at all times where the child is looking is what the child is considering as the referent. Eye movements in visual selection tasks reflect *goal-directed behavior* and as such, studies in which reference is necessary to achieve some goal (such as acting on spoken instructions) permits a researcher to infer referential and syntactic decisions.

3.2.1 *Experimental support*

Is there evidence supporting this linking assumption? Let us return for a moment to the eye movement record illustrated above in Figure 5, which involved the utterance *Put the frog that's on the napkin into the box.* Recall that upon hearing *frog*, gaze probability was split equally between the two frogs in the scene. In contrast, upon hearing *napkin*, looks did not split between the two napkins but instead converged only on the target (the frog and the napkin underneath it). It was suggested that this eye pattern for *the napkin* reflected a particular syntactic parse that children were pursuing for the phrase *that's on the napkin*: It was parsed as a relative clause modifier of the NP *the frog*, and hence required the NP *the napkin* to refer to the napkin under the frog.

One could, however, argue that this eye movement pattern is not reflecting structural and referential decisions. For instance, it could be a reflection of a simple conjunction heuristic: The child has heard *frog* and *napkin* and hence he/she looks to the only quadrant that contains both a frog and a napkin.

There are, however, several ways to design a study that would rule out this possibility and lend further support to the assumption that eye patterns are reflecting the referential implications of parsing choices. For instance, the Trueswell et al. (1999) study also contained target utterances like the following.

(1) *Put the frog on the napkin into the box.*

The absence of the *that's* in this sentence makes *on the napkin* temporarily ambiguous between being a modifier of the NP *the frog* (i.e., a property of a frog) or a goal of the verb *put* (i.e., where to put a frog). Essentially all theories of human sentence processing predict that listeners should initially parse this ambiguous *on the napkin* as a goal rather than a modifier (only to have to revise this parse upon hearing *into the box*). Some theories predict this preference based on lexical facts: The verb *put* tends to take a goal, usually in the form of a prepositional phrase (PP). If the child knows this fact, he/she will parse *on the napkin* as a goal. Other theories predict this goal preference on the grounds of structural simplic-

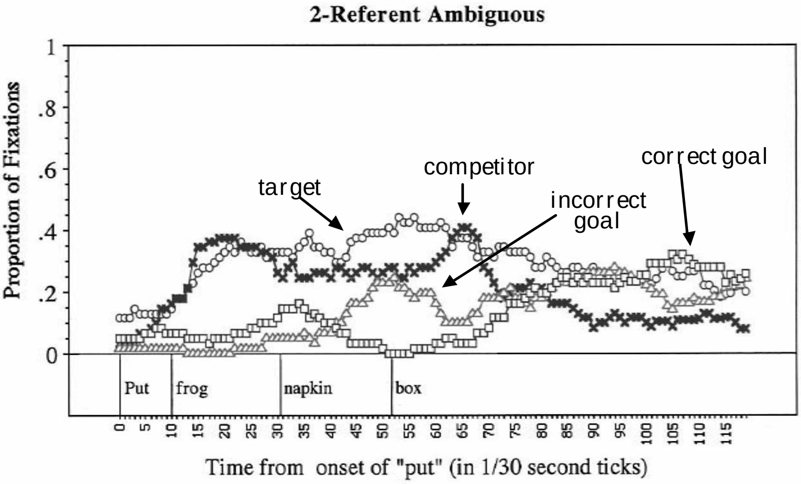


Figure 6. Children's (5-year-olds') proportion of looks over time to potential referent objects in response to *Put the frog on the napkin into the box*. From Trueswell, Sekerina, Hill and Logrip (1999). Copyright 1999, Elsevier Press.

ity; linking a PP to a verb is claimed to be computationally simpler than linking it to an NP. For either of these parsing reasons, our linking assumptions lead us to predict that children (if they parse in one of these manners) should under these conditions start considering the empty napkin (the incorrect goal) as a possible referent of *napkin*. This is because the most plausible goal for putting a frog in this case is the empty napkin (not the napkin that already has a frog on it). If a simple conjunction heuristic is at work, the result should be similar to the unambiguous sentence (i.e., we should again see increased looks to target – the only quadrant that has both a frog and a napkin). Figure 6 presents the probability plot for this temporarily ambiguous sentence.

Consistent with the parsing/reference linking assumptions, looks to the incorrect goal do in fact increase soon after hearing *napkin* in this condition, a pattern that is reliably different from that in Figure 5 when the phrase was unambiguously a modifier. Notice also that as a consequence of interpreting *on the napkin* as a goal rather than a modifier phrase, children are having trouble distinguishing between the two frogs (they are looking equally at both the target and competitor frogs for an extended period of time; see Figure 5). This additional pattern is also expected under the parsing and reference assumptions; if *on the napkin* isn't parsed as a modifier (but rather as a goal), then this phrase is no longer informative for distinguishing between the two frogs.

Since the publication of Trueswell et al. (1999) numerous other studies have been conducted that also use children's eye movement patterns during spoken language comprehension to infer ongoing syntactic and referential decisions. Snedeker and Trueswell (2004) showed that children's initial parsing preferences for syntactic ambiguities like that in example sentence (1) are the product of children's sensitivity to verb biases and not a simplicity parsing heuristic. Similar conclusions have been drawn by Kidd and Bavin (2005).

With respect to the study of reference resolution in children, several research teams have used eye movements to explore children's developing ability to resolve referential ambiguities associated with pronouns (Arnold, Brown-Schmidt, & Trueswell in press; Sekerina, Stromswold, & Hestvik 2004; Song & Fisher 2005). This work shows that children as young as 3 years of age quickly and rapidly use the gender morphology of personal pronouns (he/she) to resolve otherwise ambiguous referents. However, less reliable predictors of referent choice (such as some discourse factors related to order of mention) appear to take longer for the child to master.

Nadig and Sedivy (2002) and Epley, Morewedge and Keysar (2004) have examined children's eye movements in referential communication tasks, so as to explore the extent to which children compute the visual perspective of their interlocutors. This gaze information can in principle help determine what are plausible referents for utterances (i.e., help determine which referent objects are plausibly in common ground). Currently, there is some debate regarding children's ability to use this information; Nadig and Sedivy (2002) obtain positive results whereas Epley et al. (2004) find that children fail to use common ground in such communication tasks. (See an explanation of this apparent discrepancy below, in Section 3.2.2.) In addition, our lab has recently begun exploring how speaker gaze direction might be used by the child listener to infer parses of ambiguous strings and even the meaning of novel verbs (Nappa, Trueswell, & Gleitman 2006).

Cross-linguistic comparisons of child sentence processing abilities are also starting to use the visual world method. For instance, Choi and Trueswell (in preparation) have been exploring children's parsing preferences in Korean, a head-final language in which verbs routinely occur at the end rather than the beginning of imperative sentences. This work shows striking similarities across languages in the child's ability to use detailed lexical-syntactic / morpho-syntactic probabilities in real time, so as to estimate the most likely intended structure of a sentence. Also, Sekerina and Brooks (2007) have been exploring Russian children's word recognition abilities in various visual and linguistic contexts.

And finally, Snedeker and colleagues have been using the visual world method to understand structural priming patterns in comprehension (Snedeker & Thothathiri this volume; Thothathiri & Snedeker in press) and to explore children's understanding of quantification and scope (Huang & Snedeker 2007). These studies use processing patterns to ask questions about the underlying linguistic representations that children are forming during development. For instance, the presence of structural priming patterns that are independent of the particular lexical items used in an utterance can in principle be quite informative for issues pertaining to the levels of abstraction that children are able to operate over when acquiring and processing a language.

It is important to note that most of these recent studies used the particular method described above, in which participants act upon objects based on spoken instructions. However, some studies (e.g., Arnold et al. in press) asked children to decide if a spoken sentence accurately described a visually co-present picture. Here the goal-directed behavior also requires linking speech to the visual referent world, and as a result can provide informative patterns related to referent resolution. It has been our experience that simply asking children to passively "look at pictures and listen to a story" leads them to become more easily distracted and less likely to inspect the visual scenes. This observation is consistent with the linking assumptions discussed above: goal-directed behavior that requires referential linkage to the world is much more likely to yield interpretable eye movement patterns.

3.2.2 *Developmental interactions*

Before closing this discussion of using eye movements to infer parsing and referential decisions, it is important to explore for a moment the possibility that facts about general cognitive development might also interact with our visual world measures. For instance, the adult ability to dynamically and flexibly reconsider possible interpretations of a sentence *on the fly over the course of the sentence* no doubt requires some skill to execute in a timely manner. What general cognitive skills, if any, might be needed to achieve this? And do children have these prerequisite cognitive abilities, or do they show a protracted developmental profile?

It is well known, for instance, that for nonlinguistic tasks, children 12 years of age and younger show difficulty reinterpreting situations and inhibiting prepotent responses to stimuli (e.g., Bialystok 2001; Bunge, Dudukovic, Thomason, Vaidya, & Gabrieli 2002). For instance, children under the age of 6 years show difficulty discovering the alternative interpretation of ambiguous figures (Bialystok & Shapero 2005). Similarly, children in this age range show difficulties overriding a rule that they have recently learned for characterizing a stimulus, as in the Wisconsin Card Sorting Task, where children continue to sort based on the original rule while normal adults can switch rules with relative ease. (For discussion of these and related experimental findings see Davidson, Amso, Anderson, & Diamond 2006, and references therein.)

Put another way, children are 'cognitively impulsive.' Automatic and/or highly learned responses to stimuli are often difficult for a child to rescind and revise. Observed changes in this behavioral pattern over development follow nicely from what is known about the brain systems that support 'cognitive control' of this sort. Frontal lobe systems (particularly, left and right prefrontal cortex) have been implicated in adults to support a range of cognitive control abilities (e.g., Bunge et al. 2002; Thompson-Schill et al. 1998). These very same brain regions also show some of the most delayed neuroanatomical maturational profiles; for instance, myelination of neurons within these frontal systems is not complete until quite late in normal human development, i.e., as late as 5–10 years of age, if not later (see Diamond 2002, for a discussion).

Interestingly, this cognitive impulsivity was also observed for the 5-year-olds in the Trueswell et al. (1999) *Put*-study. Consider again the eye movement patterns in Figure 6, for the ambiguous sentence *Put the frog on the napkin into the box.* Children *never* consistently converged on the intended target frog (looking just as often at the competitor), suggesting that they never realized that *on the napkin* could be a Modifier of the NP *the frog*. In fact, children's ultimate actions suggested they had not fully recovered from their garden-path: Children were at chance selecting between the two frogs, and frequently (60% of the time) moved the selected frog to the incorrect goal – placing the frog on the empty napkin, or placing the frog on the empty napkin and then into the box. This difficulty was clearly related to ambiguity, since these same children made essentially no errors in response to unambiguous materials: *Put the frog that's on the napkin into the box.*

Researchers who are not predisposed to thinking of child language use as an emerging dynamic process might interpret such child failures as indicating an age range at which children lack some knowledge; perhaps they have not yet 'acquired' the restrictive (NP modifying) PP structure. Two findings rule out this possibility and point to the account of children 'failing to revise'. First, Hurewitz et al. (2001) showed that children within the relevant age range can produce restrictive NP modifiers (e.g., *the frog on the napkin, the one on the napkin*) when asked about two visually co-present referents. These same children nevertheless go on to misinterpret *put*-sentences like (1) in exactly the same way as that found by Trueswell et al. (1999). This suggests that in the traditional sense, these children have acquired this structure.

Second, similar parsing failures in comprehension have recently been seen in a special population of adults – specifically, an individual with a focal lesion to frontal lobe regions known to be responsible for cognitive control (Novick 2005; see also Novick, Trueswell, & Thompson-Schill 2005, for a discussion). In this study, this individual (Patient NJ) was given a battery of neurocognitive tasks designed to test frontal lobe functioning, but also the *Put*-task described above. NJ showed the characteristic deficits in cognitive control, i.e., not being able to inhibit pre-potent responses to various stimuli. Interestingly, NJ showed a parsing pattern quite similar to 5-year-olds; he failed to revise on ambiguous trials, moving (for instance) the frog onto the empty napkin (the incorrect goal) and then into the box (the correct goal). Like 5-year-olds, NJ made no errors on unambiguous versions of these sentences (*Put the frog that's on the napkin into the box*). Also quite interestingly, NJ has been found to have difficulty resolving highly biased ambiguous words as well (Novick, Bedny, Kan, Trueswell, & Thompson-Schill in preparation). Thus, NJ, who has deficits in cognitive control, shows precisely the sorts of linguistic processing deficits one might expect if cognitive control plays a role in parsing and interpretation, i.e., an inability to recover the subordinate meaning of a highly biased ambiguous structure or a highly biased ambiguous word.

This surprising association between specific frontal lobe deficits and garden-path recovery bodes well for dynamic processing accounts of child language development. Given that frontal-lobe neural systems are some of the last regions of the brain to fully mature anatomically, it is completely plausible that children's dynamic processing systems are hindered by delayed development of systems responsible for engaging cognitive control, specifically the ability to recharacterize otherwise supported interpretations of linguistic input. Khanna, Boland and Cortese (2006) explore this and related hypotheses as they pertain to children's developing ability to resolve word sense ambiguity.

Interestingly, children's inability to use joint-attention contextual constraints in referential communication tasks may be related to these issues. For example, Epley et al. (2004) found that 5-year-olds act egocentrically when selecting a referent, sometimes picking as a referent an object that was visible only to the child and not to the adult speaker. However, picking the intended "common-ground" object (i.e., the object that was visible to both the speaker and the listener) always required the child to select the subordinate meaning (or less prototypical meaning) of the referential expression. When this is controlled for (as was the case in Nadig & Sedivy 2002), children's use of joint eye gaze returns. Taken together, the data suggest that children weigh multiple linguistic and nonlinguistic constraints when making referential decisions. However, if this multiple constraint process requires overriding potent linguistic tendencies, cognitive control is necessary and difficulty may ensue.

4. Summary and conclusions

A common way of thinking about cognitive development is as the gradual acquisition of knowledge about the world. Alternatively, we can think of development as the acquisition of dynamic skills: We learn *how to interact* with the world; we learn *how to produce and comprehend* a language, etc. This chapter has reviewed some key experimental findings within the developmental psycholinguistics literature that encourages this dynamic, functional way of thinking about language learning. Studies that have recorded the eye movements of young children as they hear spoken instructions have to date been quite successful at uncovering ongoing referential and syntactic processes as they occur over the course of hearing each sentence.

An evaluation of the linking assumptions necessary to interpret findings from this methodology suggests that these assumptions are valid. However, caution and care is necessary when performing such research because developmental changes in attentional control and cognitive control can in principle interact with observations from this method. It is important to note that this concern is true of any experimental method when applied to the study of development; the onus falls on the developmental researcher to understand and even seek out these interactions in their experimental findings. Otherwise, developmental observations can be easily misattributed to the researcher's theoretical topic of interest. For instance, the present evaluation of the visual world methodology suggests that care must be taken in understanding how general attentional control and cognitive control change with age. Developmental shifts were identified in the relative contribution of exogenous and endogenous factors when it comes to the direction of spatial attention, particularly in younger children (3 years of age and younger). In addition, developmental shifts exist in general cognitive control abilities well into a child's 10th year of life. Children show a domain-general difficulty overriding initial characterizations of stimuli. This same difficulty is also manifested in language processing: Children sometimes have difficulty overriding their initial characterization of a sentence and hence sometimes fail to recover from garden-paths.

No doubt as our understanding of visual attention in the infant and child grows, significant advances will also simultaneously occur in our understanding of language learning and language processing, particularly in the relatively natural setting of discussing visually co-present referents. The visual world method serves as an important new way of evaluating the dynamics of language use in the young child. Significant theoretical advances have been made through the application of this and other real-time measures of language use. Indeed, the visual world method in particular has shown itself to be extremely valuable for understanding language representation and use as it develops from infancy into adulthood. The method is well suited for experimental investigations at multiple levels of linguistic representation (phonological, lexical, syntactic and referential) and offers important insight into the fine-grain temporal dynamics of these systems as they grow and mature.

Acknowledgments

This work was partially funded by a grant from the National Institutes of Health (1-R01-HD37507).

References

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition: Evidence for continuous mapping models. *Journal of Memory and Language, 38,* 419–439.

Arnold, J. E., Brown-Schmidt, S., & Trueswell, J. C. (in press). Pronoun comprehension in young children. *Language and Cognitive Processes.*

Aslin, R. N. & McMurray, B. (2004). Automated corneal-reflection eye-tracking in infancy: Methodological developments and applications to cognition. *Infancy, 6*, 155–163.

Ballard, D., Hayhoe, M., & Pelz, J. (1995). Memory representations in natural tasks. *Cognitive Neuroscience, 7*, 66–80.

Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition.* New York NY: CUP.

Bialystok, E. & Shapero, D. (2005). Ambiguous benefits: The effect of bilingualism on reversing ambiguous figures. *Developmental Science, 8,* 595–604.

Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J., & Gabrieli, J. D. E. (2002). Development of frontal lobe contributions to cognitive control in children: Evidence from fMRI. *Neuron, 33*, 301–311.

Choi, Y. & Trueswell, J. C. (in preparation). Putting first things last: The Korean kindergarten-path effect. Manuscript in preparation.

Cohen, M. E. & Ross, L. E. (1977). Saccade latency in children and adults: Effects of warning interval and target eccentricity. *Journal of Experimental Child Psychology, 23,* 539–549.

Cohen, M. E. & Ross, L. E. (1978). Latency and accuracy characteristics of saccades and corrective saccades in children and adults. *Journal of Experimental Child Psychology, 26*, 517–525.

Colombo, J. (2001). The development of visual attention in infancy. *Annual Review of Psychology, 52*, 337–367.

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language. *Cognitive Psychology, 6*, 84–107.

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4–13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia, 44,* 2037–2078.

Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry. In D. Stuss & R. Knight (Eds.), *Principles of frontal lobe functioning* (pp. 466–503). New York NY: OUP.

Epley, N., Morewedge, C. K., & Keysar, B. (2004). Perspective taking in children and adults: Equivalent egocentrism but differential correction. *Journal of Experimental Social Psychology, 40,* 760–768.

Fernald, A. & Hurtado, N. (2006). Names in frames: Infants interpret words in sentence frames faster than words in isolation. *Developmental Science, 9*, F33–F40.

Findlay, J. M. (2004). Eye scanning and visual search. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 135–160). New York NY: Psychology Press.

Franconeri, S. L. & Simons, D. J. (2003). Moving and looming stimuli capture attention. *Perception & Psychophysics, 65*, 999–1010.

Fukushima, J., Hatta, T., & Fukushima, K. (2000). Development of voluntary control of saccadic eye movements. I: Age-related changes in normal children. *Brain Development, 22*, 173–180.

Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language, 14,* 23–45.

Guitton, D., Buchtel, H. A., & Douglas, R. M. (1985). Frontal lobe lesions in man cause difficulties in suppressing reflexive glances and in generating goal-directed saccades. *Experimental Brain Research, 58*, 455–472.

Haith, M. M., Hazan, C., & Goodman, G. S. (1988). Expectation and anticipation of dynamic visual events by 3.5-month-old babies. *Child Development, 59*, 467–479.

Hallett, P. E. (1978). Primary and secondary saccades to goals defined by instructions. *Vision Research, 18*, 1279–1296.

Henderson, J. M. & Ferreira, F. (Eds.). (2004). *The interface of language, vision, and action: Eye movements and the visual world.* New York NY: Psychology Press.

Huang, Y. & Snedeker, J. (2007). From meaning to inference: Evidence for the distinction between lexical semantics and scalar implicature in online processing and development. Manuscript submitted for publication.

Hurewitz, F., Brown-Schmidt, S., Thorpe, K., Gleitman, L. R., & Trueswell, J. C. (2001). One frog, two frog, red frog, blue frog: Factors affecting children's syntactic choices in production and comprehension. *Journal of Psycholinguistic Research, 29*, 597–626.

Johnson, M. H. (1995). The inhibition of automatic saccades in early infancy. *Developmental Psychobiology, 28,* 281–291.

Johnson, S. P., Slemmer, J. A., & Amso, D. (2004). Where infants look determines how they see: Eye movements and object perception performance in 3-month-olds. *Infancy, 6*, 185–201.

Jonides, J. & Yantis, S. (1988). Uniqueness of abrupt onset in capturing attention. *Perception and Psychophysics*, *43*, 346–354.

Khanna, M. M. Boland, J. E., & Cortese, M. J. (2006). Development of sentence context use: When and how do children know that tag is a label and not a game? Paper presented at the workshop on *On-Line Methods in Children's Language Processing.* New York NY. March 21–22, 2006.

Kidd, E. & Bavin, E. L. (2005). Lexical and referential cues to sentence interpretation: An investigation of children's interpretations of ambiguous sentences. *Journal of Child Language, 32*, 855–876.

Kowler, E. (1995). Eye movements. In S. M. Kosslyn & D. N. Osherson (Eds.), *An Invitation to Cognitive Science,* Vol. 2: *Visual Cognition* (2nd edn., pp. 215–265). Cambridge MA: The MIT Press.

Liversedge, S. & Findlay, J. (2001). Saccadic eye movements and cognition. *Trends in Cognitive Science, 4*, 6–14.

Matin, E., Shao, K. C., & Boff, K. R. (1993) Saccadic overhead: Information processing time with and without saccades. *Perception & Psychophysics, 53*, 372–380.

McMurray, B. & Aslin, R. N. (2004). Anticipatory eye movements reveal infants' auditory and visual categories. *Infancy,* 6, 203–229.

Nadig, A. & Sedivy, J. C. (2002). Evidence of perspective-taking constraints in children's on–line reference resolution. *Psychological Science, 13*, 329–336

Nappa, R., Trueswell, J. C., & Gleitman, L. R. (2006). Effects of attention and intention on sentence parsing. Paper presented at the *31st Annual Boston University Conference on Language Development*, Boston, MA.

Novick, J. M. (2005). Cognitive control and the role of Broca's area in sentence processing. PhD dissertation, University of Pennsylvania.

Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2005). Cognitive control and parsing: Re-examining the role of Broca's area in sentence comprehension. *Journal of Cognitive, Affective, and Behavioral Neuroscience, 5,* 263–281.

Novick, J. M., Bedny, M., Kan, I., Trueswell, J. C., & Thompson-Schill, S. L. (in preparation). The role of LIFG in cognitive control, language production, and language comprehension: A single-case study. Manuscript in preparation.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*, 3–25.

Pylyshyn, Z. W. (1994). Some primitive mechanisms of spatial attention. *Cognition,* 50, 363–384.

Rao, R. P. N., Zelinsky, G. J., Hayhoe, M. M., & Ballard , D. H. (2001). Eye movements in iconic visual search. *Vision Research, 42*, 1447–1463.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372–422.

Ross, S. M. & Ross, L. E. (1983). The effects of onset and offset warning and post-target stimuli on the saccade latency of children and adults. *Journal of Experimental Child Psychology, 36,* 340–355.

Ruz, M. & Lupiáñez, J. (2002). A review of attentional capture: On its automaticity and sensitivity to endogenous control. *Psicológica, 23,* 283–309.

Salman, M. S., Sharpe, J. A., Eizenman, M., Lillakas, L., Westall, C., & To, T. et al. (2006). Saccades in children. *Vision Research, 46*(8–9), 1432–1439.

Scerif, G., Karmiloff-Smith, A., Campos, R., Elsabbagh, M., Driver, J., & Cornish, K. (2005). To look or not to look? Typical and atypical development of oculomotor control. *Journal of Cognitive Neuroscience, 17*, 591–604.

Sekerina, I. A. & Brooks, P. J. (2007). Eye movements during spoken-word recognition in Russian children. *Journal of Experimental Child Psychology, 98,* 20–45.

Sekerina, I. A., Stromswold, K., & Hestvik, A. (2004). How do adults and children process referentially ambiguous pronouns? *Journal of Child Language, 31*, 123–152.

Snedeker, J. & Trueswell, J. C. (2004). The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology, 49,* 238–299.

Song, H., & Fisher, C. (2005). Who's "she"? Discourse structure influences preschoolers' pronoun interpretation. *Journal of Memory & Language, 52,* 29–57.

Spivey, M. J., Tanenhaus, M. K., Eberhard, K. M., & Sedivy, J. C. (2002). Eye movements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution. *Cognitive Psychology, 45*, 447–481.

Swingley, D., Pinto, J. P., & Fernald, A. (1999). Continuous processing in word recognition at 24 months. *Cognition, 71*, 73–108.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*, 1632–1634.

Tanenhaus, M. K. & Trueswell, J. C. (2005). Eye movements as tool for bridging the language-as-product and language-as-action traditions. In J. C. Trueswell & M. K. Tanenhaus (Eds.), *Approaches to studying world-situated language use: Bridging the language-as-product and language-as-action traditions.* Cambridge MA: The MIT Press.

Thompson-Schill, S. L., Swick, D., Farah, M. J., D'Esposito, M., Kan, I. P., & Knight, R. T. (1998). Verb generation in patients with focal frontal lesions: A neuropsychological test of neuroimaging findings. *Proceedings of the National Academy of Science, 95*, 15855–15860.

Thothathiri, M. & Snedeker, J. (in press). Syntactic priming during language comprehension in three- and four-year-old children. *Journal of Memory and Language.*

Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition, 73*, 89–134.

Trueswell, J. C. & Tanenhaus, M. K. (Eds.). (2005). *Approaches to studying world-situated language use: Bridging the language-as-product and language-as-action traditions.* Cambridge MA: The MIT Press.

Yang, Q., Bucci, M. P., & Kapoula, Z. (2002). The latency of saccades, vergence, and combined eye movements in children and adults. *Investigative Ophthalmology & Visual Science, 43,* 2939–2949.

Yantis, S. (1993). Stimulus-driven attentional capture and attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 676–681.

Yantis, S. & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance, 10*, 601–621.

Yantis, S. & Jonides, J. (1990). Abrupt visual onsets and selective attention: Voluntary verses automatic allocation. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 121–134.

Yantis, S. & Jonides, J. (1996). Attentional capture by abrupt onsets: New perceptual objects or visual masking? *Journal of Experimental Psychology: Human Perception and Performance, 22,* 1505–1513.

Zelinsky, G. J., Rao, R. P. N., Hayhoe, M. M., & Ballard, D. H. (1997). Eye movements reveal the spatiotemporal dynamics of visual search. *Psychological Science, 8*, 448–453.

CHAPTER 4

Looking while listening

Using eye movements to monitor spoken language comprehension by infants and young children

Anne Fernald, Renate Zangl, Ana Luz Portillo
and Virginia A. Marchman

The "looking-while-listening" methodology uses real-time measures of the time course of young children's gaze patterns in response to speech. This procedure is low in task demands and does not require automated eyetracking technology, similar to "preferential-looking" procedures. However, the looking-while-listening methodology differs critically from preferential-looking procedures in the methods used for data reduction and analysis, yielding high-resolution measures of speech processing from moment to moment, rather than relying on summary measures of looking preference. Because children's gaze patterns are time-locked to speech and coded frame-by-frame, each 5-min experiment response latencies can be coded with millisecond precision on multiple trials over multiple items, based on data from thousands of frames in each experiment. The meticulous procedures required in the collection, reduction, and multiple levels of analysis of such detailed data are demanding, but well worth the effort, revealing a dynamic and nuanced picture of young children's developing skill in finding meaning in spoken language.

1. Introduction

Developmental studies of comprehension in very young children have relied traditionally on off-line measures, responses made after the offset of the speech stimulus that do not tap into the real-time properties of spoken language. Studies of incremental processing by adults rely on on-line measures that monitor the time course of the listener's response in relation to key points in the speech signal. Because comprehension occurs rapidly and automatically without time for reflection, it is revealing to study the listener's interpretation *during* speech processing and not just afterward. Classic on-line behavioral techniques used to investigate incremental speech processing by adults include phoneme monitoring (Cutler & Foss 1977), gating (Grosjean 1985), and cross-modal priming (Marslen-Wilson & Zwitserlood 1989), among others. Although some of these have been adapted for use with school-aged children (Cutler & Swinney 1987; Walley 1993; Clahsen this volume), the task demands are often problematic for younger children. Research with adults using automated eyetracking techniques has been extremely productive in recent years, providing sensitive on-line measures high in ecological validity (Tanenhaus, Magnuson, Dahan & Chambers 2000). This technique also been used with 4–5-year-old children (Trueswell, Sekerina, Hill, & Logrip 1999; Snedeker & Trueswell 2004).

Here we describe a simpler but equally powerful experimental method for monitoring the time course of comprehension by infants and very young children, which we call the *looking-while-listening* procedure. In this procedure, children look at pairs of pictures while listening to speech naming one of the pictures. Their gaze patterns are videotaped and eye movements are measured with high precision in relation to relevant points in the speech signal. Using the looking-while-listening paradigm, we have shown that speed and efficiency in infants' on-line responses to familiar words in continuous speech increase dramatically over the second year for both English- and Spanish-learning children (Fernald, Pinto, Swingley, Weinberg, & McRoberts 1998; Hurtado, Marchman, & Fernald 2007), that 24-month-olds are able to process phonetic information incrementally (Swingley, Pinto, & Fernald 1999; Fernald, Swingley, & Pinto 2001), and that individual differences in children's speech processing efficiency are related to their level of lexical and grammatical development (Fernald, Perfors, & Marchman 2006). These and other recent findings validate the looking-while-listening paradigm as a powerful new method for exploring how very young children take advantage of perceptual and linguistic features of the speech they hear in relation to information in the visual world, as they learn to find meaning in spoken language.

1.1 Research on the early development of receptive language skills

Long before they can speak their first words, infants begin to reveal their emerging knowledge of language by responding meaningfully to the speech they hear. Yet because comprehension is a mental event that can only be inferred indirectly from a child's behavior in a particular context, the early development of receptive language competence has been less accessible to scientific inquiry than developmental gains in speech production. However, research on infant cognition over the past three decades has led to valuable experimental methods for "reading the minds" of very young language learners, techniques that have made it possible to explore the developmental origins of understanding in greater depth and with greater precision. Much of this research has examined how infants become attuned to sound patterns in the ambient language over the first year of life (e.g., Werker & Tees 1984; Kuhl 2004), and attend to speech patterns relevant to language structure (e.g., Jusczyk 1997; Saffran 2002). These studies show that over the first year infants become skilled listeners, able to make distributional analyses of phonetic features of spoken language, and that they form some kind of acoustic-phonetic representation for frequently heard sound patterns (e.g., Halle & de Boysson-Bardies 1994). Such accomplishments are often cited as evidence for early "word recognition". But since this selective response to familiar words can occur without any association between particular sound patterns and meanings, it is perhaps appropriately viewed as evidence for pattern-detection abilities prerequisite for recognizing words in continuous speech. Much less is known about how language-specific processing strategies continue to develop beyond the first year, as children in their second and third years begin to appreciate regularities at higher levels of linguistic organization, using their emerging lexical and morphosyntactic knowledge to make sense of spoken words and sequences of words in combination (see Fernald & Frank in press; Naigles 2002).

1.2 Observational and off-line experimental measures of children's skill in comprehension

Between the ages of 10 and 14 months children typically begin to show signs of associating sound patterns with meanings, speaking a few words and appearing to understand many more. By the end of the second year, they reveal progress in understanding through increasingly differentiated verbal and behavioral responses to speech. But growth in receptive language competence is harder to observe than growth in productive abilities, because the processes involved in comprehension are only partially and inconsistently apparent through the child's spontaneous behavior. Scientific studies of early comprehension have made use of quite different methodologies that fall into four main categories (see Fernald 2002): *Diary studies* provided the first systematic observational data on early comprehension abilities, describing how young children appear to interpret speech in their everyday activities and interactions (e.g., Lewis 1936; Bloom 1973). *Studies of vocabulary growth* use parental-report checklists to track changes in the estimated size of the child's receptive lexicon; such changes correlate in interesting ways with later grammatical development (e.g., Fenson, Marchman, Thal, Dale, Reznick, & Bates 2007). *Naturalistic experiments on comprehension* use behavioral responses to test infants' ability to identify familiar words (e.g., Benedict 1979) and understand words in combination (e.g., Shipley, Smith, & Gleitman 1969). *Experiments on word learning* constitute the largest area of research related to early understanding: one approach focuses on how cognitive biases guide inferences about word meanings (e.g., Markman 1989); others explore how children use linguistic (e.g., Katz, Baker, & McNamara 1974) and pragmatic knowledge (e.g., Tomasello 2000) to guide these inferences.

All of these approaches rely on off-line measures of comprehension, i.e., assessments of understanding based on observing the child's behavior after hearing a particular linguistic stimulus. In the case of diary observations and parental-report checklists, the judgment that a child does or does not "understand" a word such as *dog* or *cup* is made informally by adults who interact regularly with the child in many different contexts. In the case of off-line measures used experimentally, these judgments are based on the child's behavior in a more controlled situation, with a clearly defined response measure such as choosing an object or pointing to a picture given two or more alternatives. What these measures have in common is that they all are based on children's responses to a spoken word or sentence *after* it is complete rather than as it is heard and processed. While such off-line procedures enable researchers to assess whether or not a child responds systematically in a way that indicates understanding, they do not tap into the real-time properties of spoken language and thus reveal less about the child's developing efficiency in identifying and interpreting familiar words in continuous speech.

1.3 The development of preferential-looking measures for assessing comprehension

In 1963 Robert Fantz published the first study using a preferential-looking method with infants, showing that even newborns looked selectively at some visual stimuli over others. Although Haith (1980) later questioned whether "preference" was the appropriate way to characterize infants' selective looking behavior, the findings that emerged from dozens of preferential-looking studies in this period suggested that certain early visual biases appeared to be independent of previous experience with particular stimuli. Adapting the preferential-looking procedure to investigate cross-modal perception in infants, Spelke (1976) presented infants with two visual stimuli, only one of which matched a simultaneously presented auditory stimulus, and found that infants looked significantly longer to the matching than to the non-matching visual stimulus. This auditory-visual matching procedure was later modified in different ways by investigators interested in the development of language comprehension in the early years of life.

The first experimental procedures for testing infants' knowledge of object words in a controlled setting were also introduced at the end of the 1970's. Benedict (1979) found that 12-month-olds would orient reliably to a familiar object when it was named, even when nonverbal behaviors such as gaze and pointing by the speaker were eliminated. In a more precisely controlled procedure using eye movements as an index of word recognition, Thomas, Campos, Shucard, Ramsey and Shucard (1981) compared the ability of 11- and 13-month-old infants to identify a familiar named object from an array of competitors matched for visual salience. The finding that some 12-month-olds could identify the correct referents of a few familiar words was hardly surprising, since that much was known from observational studies and parental report. What this new method offered was a way to assess word recognition more objectively. Unlike the informal observations of comprehension used earlier, this procedure enabled the experimenter to standardize stimulus presentations, to define carefully which behaviors counted as a correct response, and to eliminate gestural and other nonverbal cues from the experimenter that might indicate the target object.

The innovative studies by Spelke (1976) and Thomas et al. (1981) provided a foundation for later research by Golinkoff, Hirsh-Pasek, Cauley and Gordon (1987) and by Reznick (1990) in which preferential-looking measures were further adapted for use in assessing early language comprehension. The version of the method developed by Golinkoff et al., now known as the "intermodal preferential looking paradigm" (IPLP), has been used in numerous studies (e.g., Hollich, Hirsh-Pasek, & Golinkoff 2000; Shafer & Plunkett 1998; Meints, Plunkett, & Harris 2002). In the procedures used in these studies, infants are shown two pictures of objects as they hear speech naming one of the objects. In some studies gaze patterns are coded in real-time using a button box to record fixations on the target and distracter objects and shifts between the two pictures. The dependent measures typically used as an index of comprehension in such studies include total looking time to the target picture and duration of longest look to the target picture. With these measures, researchers using the IPLP have been able to investigate several interesting questions about children's early lexical and syntactic knowledge – for example, whether or not infants at a particular age look relatively longer to the correct target picture than to the distracter picture when asked to identify the target object in a sentence such as *Find the cat* (e.g., Golinkoff et al. 1987), or to interpret a spatial proposition in a sentence such as *Look at the cat on the table* (Meints et al. 2002).

1.4 The evolution of the looking-while-listening procedure: Moving to real-time measures

Informed by the studies of Thomas et al. (1981), Golinkoff et al. (1987), and Reznick (1990), our research group began to use a modified version of the preferential-looking method to investigate whether particular features of infant-directed speech, such as exaggerated pitch and vowel lengthening, might make it easier for young language learners to identify familiar words in fluent speech. Our initial goal in modifying the preferential-looking paradigm was to increase the sensitivity, reliability, and validity of the measures, by making minor modifications to the procedure that served to eliminate confounding variables. Earlier preferential-looking studies had used different stimuli as target and distracter objects, thus potentially confounding object salience with target status. And some had also failed to counterbalance side of target object presentation, another confound that made it difficult to interpret infants' selective looking behavior unambiguously. To address these concerns, we made sure that all target objects were also presented as distracters to reduce the influence of initial object preferences, and that side of presentation was fully counterbalanced to control for possible side bias. These and other minor but potentially influential changes were undertaken to increase internal validity in our experimental designs. In addition, we made a major change in the measures used to capture infants' gaze patterns in response to speech: Rather than coding eye movements in real-time using a button box, as was the practice in other preferential-looking procedures at the time, we began to code eye movements from the videotapes, frame-by-frame in slow motion. This change enabled us to eliminate from our measurements the noise introduced by the ca. 300 ms latency of the observer to press the button, a critical first step in the direction of achieving greater precision in our dependent measures. These modifications resulted in a labor-intensive version of the original procedure, requiring several hours to code the record of each 5-min test session, but the enhanced reliability of the measurements justified the effort.

However, we were still not yet taking full advantage of the potential for increased temporal resolution in our measures of infants' gaze patterns. In the Golinkoff et al. (1987) paradigm, word recognition was operationalized as a tendency to look longer at the named target picture than at the distracter picture, with looking-time-to-target averaged over a 6-s measurement window following offset of the speech stimulus. A similar criterion is common today in most labs using this procedure (Hollich et al. 2000; Schafer & Plunkett 1998), and we adopted it initially as well. The 6-s response window reflected a premise that seemed entirely reasonable at the time. Although psycholinguistic research with adults had recently shown that experienced listeners can process speech incrementally, generating hypotheses about word identity based on what they have heard up to that moment (Marslen-Wilson & Zwitserlood 1989), we simply assumed that infants would be considerably slower than adults in processing continuous speech. And since there were no data yet available on the time course of real-time speech processing by infants, we had to discover the hard way, through a long process of trial and error, how wrong we were in this assumption!

In three early studies we tested infants at different ages to investigate the influence of prosodic features on their ability to recognize familiar words. While these results were promising (Fernald, McRoberts, & Herrera 1992; see Fernald, McRoberts, & Swingley 2001, for a review of these findings), it became increasingly clear that we needed to understand better how the time course of infants' responses changed with development over the second year. For example, when using a percent-correct measure averaged over a 6-s measurement window, we were surprised to find that 24-month-olds apparently performed *less* well than 18-month-olds. This counterintuitive result suggested that our measures of looking time were failing to capture the gains in accuracy expected toward the end of the second year, an observation that led us to make parametric reductions in the time window over which looking-time was averaged. When we used a 4-s measurement window, the data began to make more sense, and with a 2-s window, the predicted improvement in word recognition finally became clear: for 18-month-olds the mean percentage of looking time to the target picture during this much shorter window was ca. 60%, and for 24-month-olds it rose to 80%. By adopting the 6-s coding window that was standard at the time, we had greatly underestimated the accuracy of 24-month-olds in this word recognition task; these older infants had in fact oriented quickly to the target picture upon hearing it named and had looked at it for 2–3 seconds, but then they tended to look back at the other picture or to look away. Since these look-backs and look-aways by the 24-month-olds in most cases followed a correct response within a few seconds, they were actually a sign of *rapid* processing; however, when this post-response "noise" was averaged into the percent-looking-to-target averaged over a 6-s window, the 24-month-olds looked less overall and thus appeared to be less accurate than 18-month-olds. This discrepancy was obviously not because the older children were slower or less reliable in recognizing familiar words, but rather because they responded more quickly than the younger children and then tended to lose interest in the target picture midway through the 6-s window.

It was clear from these analyses that infants were able to identify familiar spoken words much more rapidly than we had imagined possible. Thus at this point we made two important procedural changes: First, we began to measure infants' eye movements from the *onset* of the target word, not the offset. And second, rather than averaging looking time over an arbitrary, fixed coding window (which might be suitable at one age but not at another), we began to code eye movements at the finest level of resolution possible given the limits of our technology. In the first published study using time course measures of spoken word recognition with infants (Fernald et al. 1998), this limit was 100 ms, the resolution of our time-code generator at that time. In subsequent studies the level of resolution increased to 33 ms, the duration of a single video frame. This change was critical in enabling us to move from the global measure of total looking time to the target picture to a much more precise measure of reaction time, capturing the child's latency to shift from the distracter to the target picture. Through this incremental process of refining our analysis techniques and improving our coding technology, the looking-while-listening procedure has become an increasingly powerful method for monitoring the time course of infants' comprehension of continuous speech, enabling us to measure reaction time as well as accuracy in word recognition in very young language learners. The result is a testing procedure with minimal task demands that can be used with both infants and adults, yielding eye movement data comparable in reliability and precision to data from adult studies that require technically more sophisticated automated eyetracking methods (e.g., Tanenhaus et al. 2000; Henderson & Ferreira 2004).

2. From preference to reference: A nuts-and-bolts overview of the looking-while-listening paradigm

In this section we provide an overview of the looking-while-listening procedure, following the traditional format of a Methods section in a research report (i.e., describing participants, procedure, etc.). However, consistent with the focus on methodology in this volume, the information in each section extends beyond the details relevant to any particular study, integrating our experience using this paradigm with many different experimental designs and with participants ranging from 12-month-old infants to adults. Our goal here is to present an overview of the looking-while-listening paradigm at a functional level, discussing the logic of each step in the procedure, from preparing and running an experiment to coding eye movements and analyzing the data using several different measures of efficiency in spoken language processing.

As mentioned in the introduction, the looking-while-listening procedure is superficially similar to a preferential-looking procedure in that infants are shown two pictures on each trial and hear speech naming one of the pictures as we record their gaze patterns in response to the speech signal. But the static notion of "preference" is irrelevant for our purposes. Rather than construing infants' looking behavior in response to spoken language as motivated by *preference*, we are interested in how children establish *reference* by making sense of spoken language from moment to moment, a process of incremental interpretation that is highly dynamic in adult comprehension (e.g., Knoeferle & Crocker 2006). Rather than relying exclusively on a single preference score based on total looking to the target averaged over a fixed coding window following the speech stimulus, we are also interested in the time course of looking to the referent as the sentence unfolds. To achieve this goal we code children's gaze patterns off-line through careful frame-by-frame inspection, enabling precise measurement of their latency to initiate an eye movement toward the appropriate referent, time-locked to critical moments in the speech signal on each trial. Thus the looking-while-listening procedure incorporates the same sensitive temporal measures used in eyetracking studies with adults and older children (see Trueswell & Tanenhaus 2005), differing from those methods in only three noteworthy respects: first, we typically use visual displays with only two alternatives rather than more complex scenes involving four or more displays; second, we do not to use an automated eyetracker, because comparable precision and reliability can be achieved using our high-resolution video coding procedures; and third, an inherent constraint in working with infants and very young children is that they have limited and fluctuating attention; thus we have to design experiments lasting just a few minutes with only 30–40 trials, which yield much less (and potentially noisier) data than the longer experiments with hundreds of trials that are possible in studies with adults.

2.1 Participants

The looking-while-listening paradigm has been used effectively in our laboratory with infants as young as 14 months of age as well as older children and adults. When we tested 12-month-olds in several pilot studies using this procedure, we found their performance to be close to chance, as did Zangl, Klarman, Thal, Fernald and Bates (2005) in a study using a similar procedure. While 12-month-olds are happy to look at the pictures displayed, they are likely to fixate on only one picture and ignore the other on any given trial, shifting less frequently overall than infants just two months older. This pattern of results with younger infants reflects their limited linguistic knowledge, but may also result from attentional limitations at this age.

2.2 Preparing visual stimuli

In studies in which we assess recognition of familiar words, the visual stimuli consist of pictures of real objects with which infants and toddlers are very familiar. The target objects chosen for each study have names that are highly likely to be understood by children in the relevant age range, based on vocabulary norms for the MacArthur-Bates Communicative Development Inventory (Fenson et al. 2007). Realistic images of common objects judged to be prototypical for children at each age are drawn from various sources, including image banks on the Internet and pictures taken with digital cameras. For studies in which children are taught and then tested on novel words, we use pictures of constructed objects with which children have no prior experience. In some cases, objects are shown against a uniform gray background; however, in studies where the same target objects are used repeatedly across trials, objects may be presented on more diverse and complex backgrounds that vary in color and pattern, in order to maintain children's interest.

All images used as stimuli are edited so the objects are approximately the same size and are informally matched in visual complexity and brightness. When choosing and editing pictures for a particular study, it is important to balance the visual salience of the target and distracter pictures on each trial, keeping in mind that any given picture is not inherently "salient" but only in relation to other pictures in the stimulus set. If one picture in a pair is much more engaging than another, it is more difficult to tell whether looking to the target picture on a given trial was influenced by the speech stimulus rather than by baseline visual preferences. We have found that younger infants in particular generally find images of animate objects more interesting than inanimate ones, and thus might initially attend more to a picture of an animal if it was paired with a picture of an artifact. To reduce the effects of this potential bias, we present yoked pairs of animates (dog and baby) on some trials, and pairs of inanimates (shoe and car) on others. Although counterbalancing measures, such as presenting each stimulus picture as both target and distracter an equal number of times in every experiment, are designed to mitigate the effects of differential visual salience among stimuli, such differences should be reduced as much as possible because they contribute noise to the data.

In some experiments the potential for visual salience differences is particularly acute due to the nature of the stimulus objects used, and in this case pilot testing is advisable. To evaluate the suitability of new images as potential stimuli, candidate pictures are paired with other potential stimulus pictures and presented to children in the appropriate age range without any verbal information, to determine whether participants orient more to one picture than the other based on differences in visual salience. If children do indeed fixate one picture significantly more than the other, the pictures need to be edited further to adjust factors like hue, saturation, and visual complexity that might contribute to differential salience. The modified stimuli then need to be pilot-tested again until children are equally likely to fixate both pictures. Balancing the relative visual salience of stimulus objects is an important step in maximizing the sensitivity of the measures in the looking-while-listening procedure, especially when working with infants younger than two years of age. This extra step is not as labor-intensive as it may seem, if such "picture-pilot" trials for potential stimuli in a new experiment can be included as filler trials in another ongoing experiment with children in the same age range.

2.3 Preparing auditory stimuli

In the looking-while-listening procedure children's eye movements are examined in relation to words that are heard at critical points in the unfolding speech stream, and so precision in recording and editing the speech stimuli is key for all further analyses. To produce the stimuli used in our studies, multiple tokens of each sentence are digitally recorded by a native speaker of English or Spanish, then acoustically analyzed, and edited using *Peak* or *Praat* sound-editing software. Because it is a goal in most of our studies to achieve comparability in relevant acoustic dimensions across the set of speech stimuli, recording ten or more tokens of each stimulus sentence yields a range of candidate stimuli to choose from. After measuring each relevant element in every recording, we then choose the tokens of each sentence type that are most closely matched in the duration of critical words as well as in overall prosodic characteristics, across the whole stimulus set. The criteria for final stimulus selection vary to some extent with the experimental question. For example, in a recent study (Zangl & Fernald 2007) investigating children's responses to grammatical and nonce determiners (*Where's the dog?* vs. *Where's po dog?*), the choice of tokens for the final stimulus set was governed by three considerations: First, all carrier frames were matched in duration (M = 269 ms; range: 261 to 278). Second, grammatical and nonce articles were matched in duration (grammatical: M = 160 ms; range: 147 to 172; nonce: M = 168 ms; range: 156 to 172). And third, it was also important that all articles be unstressed, and to the same degree. When rated by two phonetically trained listeners, both sentence types were judged to be comparable in intonation contour. In some studies, cross-splicing techniques are also used to increase stimulus control. In the Zangl and Fernald study, for example, it was important to ensure that each target noun was acoustically identical for each sentence type, and so a single token of each target noun was cross-spliced into carrier frames across conditions.

Depending on the goals and design of the study, measurements are made for particular regions of interest in each stimulus sentence, such as carrier frames, determiners, adjectives, target nouns, etc., to enable reliable identification of the acoustic onset and offsets of key words that are crucial for measures of response latency. For example, when investigating children's efficiency in interpreting familiar object names in simple English sentences such as *Find the car*, determining the onset and duration of the target word is sufficient to time-lock the child's response to the unfolding object name (e.g., Fernald et al. 1998, 2006). In this case, the noun onset is the critical point at which the child begins to accumulate phonetic information necessary to interpret the target word and identify the referent object. However, the critical point may come earlier in cases where linguistic information encountered *before* the noun is potentially informative. For example, nouns in Spanish have obligatory grammatical gender and are preceded by gender-marked determiners. In a study investigating whether Spanish-learning children could take advantage of gender-marked articles to identify an upcoming object name, Lew-Williams and Fernald (2007) measured children's eye movements to the target object in relation to the first potentially informative element in the sentence, in this case the prenominal article. This study required measurement of the onset and offset of the determiner as well as the target noun, and tokens of each stimulus sentence were chosen to maximize comparability on these dimensions across the stimulus set. That is, the best-matching tokens were selected from the larger set of exemplars recorded for each sentence. These were then further edited so that key points in the speech stream could be carefully matched within one frame across sentences.

By making the critical speech stimuli in each condition prosodically comparable, we can reduce the effects of potentially confounding factors that might influence children's responses. However, maximizing experimental control in this way also has a less desirable consequence: we end up with a stimulus set that is highly repetitive and potentially monotonous overall, which has drawbacks in terms of maintaining children's attention. Experience has shown us that repeating short sentences with the same intonation contour on trial after trial causes children to lose interest sooner than when they have a more varied sequence of speech stimuli to listen to. One step in reducing this problem is to add a tag question or statement to the speech stimulus on each trial, following the sentence containing the target word sentence by a 1-s interval. These short sentences are all attention-getters such as *Can you find it?* or *Check that out!* The same speaker who produced the stimuli for the study records these tag sentences with lively intonation that varies from token to token, and they are appended to the speech stimuli in a way that maximizes variability from trial to trial by alternating sentences with rising and falling pitch contours. A second step in increasing variability is to intersperse test trials with filler trials, as described below.

2.4 Constructing stimulus orders

The visual and auditory stimuli are combined to construct two or more different stimulus orders. Although in earlier studies using this method we included only 10–20 trials in each experimental order, in more recent studies we have learned from experience how to get high quality data on a larger number of trials. One insight was the importance of incorporating the right amount of visual and auditory complexity in experimental and filler trials across the entire stimulus set, finding the appropriate balance that is optimal for a particular age group. By fine-tuning stimulus sets in this way, we can now test 18 month-olds on 30–40 trials, and older children on 40–50 trials. Trial types are counterbalanced so that the target object appears on the left and right sides an equal number of times, and does not appear on the same side more than three trials in a row.

In addition to using tag sentences to increase prosodic variability within experimental trials, it is also important to include filler trials that introduce visual as well as acoustic novelty across the stimulus set. If a study requires using multiple tokens of trials of one particular type, then the total number of trials may include up to 30% filler trials, designed to reduce repetitiveness and to maintain children's interest. Filler trials might consist of jazzy pictures that are visually more complex and colorful than the experimental stimuli, such as a market scene or hot air balloons, accompanied by an exclamation such as *Good job – here's another picture!* They might also consist of trial types we are pilot testing for other experiments, as long as the target objects are sufficiently different from those the children are being tested on. Thus one critical factor in maximizing children's attentiveness throughout the experiment is to carefully balance variation and repetition across the stimulus set. Another critical factor is to provide a stimulus set that is challenging, but not too challenging, for children at each particular age.

2.5 Apparatus

The looking-while-listening procedure is conducted in a sound-treated room containing a test booth with two cloth-covered side walls, shown in Figure 1A. Built into the third wall of the booth is a polycarbonate rear-projection screen that is tough and scratch resistant. Behind the screen is a rear-screen projector designed with very short throw-space; thus the distance from the projector to the screen is only 1 m, ideal for use in a small room. The video camera is positioned below the screen, mounted out of reach of the child. During testing, younger children sit in the booth on the parent's lap on a swivel chair, with the seat height adjusted so that the child's eyes are at the appropriate level in relation to the video

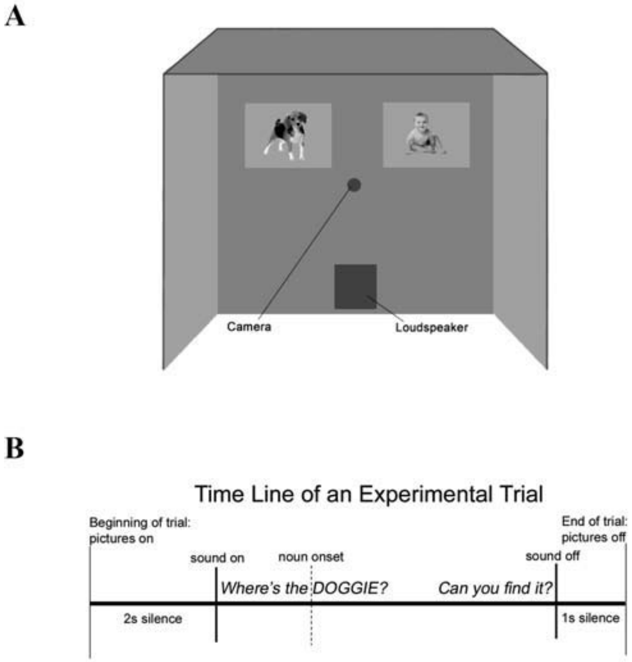


Figure 1. (A) Configuration of test booth with rear-projection screen used in the looking-while-listening procedure; (B) schematic time line for a typical trial.

camera, and the child's head is ca. 1 m from the screen. Children 4 years and older sit alone on a non-swivel chair, also adjusted to the correct height so their eyes are centered on the video recording. When a parent is holding the child during the test session, the adult wears dark glasses made opaque by covering the lenses on the inside with cloth tape, so the visual stimuli cannot be seen. Although picture size may vary from one experiment to another, the pictures used in our standard setup are projected at a size of ca. 36 × 50 cm, aligned horizontally at a distance of 60 cm from one another, with the center of each image slightly above the child's eye level. Speech stimuli are delivered through a central loudspeaker located on the floor below the screen. The black-and-white analog video camera, sensitive to low light levels, is fitted with a zoom lens that is focused on the child's face only. Positioning each child relative to the camera and focusing the camera for each session are crucial to collecting useable data; only when the child's eyes are clearly visible is it possible to reliably identify the shifts in fixation that are crucial for calculating response speed.

The experiment is run using *PsyScope* software controlled by a Macintosh computer located in an adjacent control room. On each trial this computer out-puts the visual stimuli to the rear-screen projector in the testing room, and also to a monitor in the control room. The video record of the child's eye movements on each trial is sent to a mixer in the testing room which integrates the video signal with graphic information about participant and trial number, stimulus order, and the onset and offset of auditory and visual stimuli. This combined information is then fed to a time code generator where it is time-stamped and sent to a VCR which records the complete video signal, including the child's eye movements, the event information, and the time code. This composite video record is then digitized on a second computer in the control room.

2.6 Testing procedure

Each test session is preceded by a 15-min period during which the parent completes the procedures for informed consent and the experimenter interacts with the child. When both parent and child are comfortable and ready to begin, they are escorted to the testing room. The lights in the testing room are already dimmed, and images of puppets or cartoon characters familiar to young children are on the screen. A lamp mounted on top of the screen provides just enough light to avoid complete darkness, ensuring that the visual display is the most interesting thing in sight for the child. In a typical testing situation, two experimenters are involved: the first experimenter seats the parent and child in the testing room and checks that the camera is optimally positioned; the second experimenter, located in the control room, speaks to the child over the microphone to familiarize her with the sound source and directs attention to the visual display. If necessary, the second experimenter also coordinates with the experimenter in the testing room to focus the camera on the child's face. Once the child is at ease and the parent has put on the darkened sunglasses, the first experimenter leaves the room and the test session begins.

On a typical test trial, the target and distracter pictures are shown for 2 s prior to the onset of the speech stimulus, providing the child with enough time to check out both pictures. As mentioned earlier, using each object as target in one trial and distracter in another trial constitutes an important control, offsetting the problem that one picture may be fixated first because it was named on a previous trial or simply because it is visually more salient. Figure 1B shows the time line for a typical trial. Trials last on average 6–8 s, including the 2 s before the onset of the speech stimuli, and are separated by an 800 ms brief interval when both screens are blank. The entire experiment lasts about 5–6 min.

In eyetracking studies with adults, participants are instructed to look at a central fixation point before responding to the verbal stimulus, and on most trials they willingly comply. This step has the advantage that the adult starts each trial at the same neutral location equidistant from all the visual stimuli; thus in principle a latency to shift from the center point to one of the stimulus objects can be assessed on every trial. A similar method has been used in eyetracking experiments with preschool children, who are requested to look at a central smiley face at the start of a new trial (see Trueswell this volume). In the looking-while-listening procedure we do not use a central fixation point, for the simple reason that infants and very young children will not follow such instructions. Although some preferential-looking studies have used a central fixation point at trial-onset, none have reported data on how long children actually maintain fixation at center. We explored this issue in a recent study by presenting a bright geometric image between the two stimulus pictures to 26- and 36-month-olds, in an effort to bring their attention to center at the start of the speech stimulus (Portillo, Mika, & Fernald 2006). However, within milliseconds children at both ages had shifted randomly to one side or the other *before* the target word was spoken. Since most children on most trials were already looking at the target or distracter object before hearing the target word, there is no evidence that a central fixation point is actually effective in studies with very young children.

2.7 Pre-screening for codeable trials

The video record of each test session is integrated with information about the participant (e.g., subject number, gender, age) as well as information relevant for data analysis (e.g., stimulus order and trial type) using the custom software *Eyecoder*. This information is not visible to observers during prescreening or coding. As shown in Figure 2, the video record of the test session also includes additional information that is visually accessible and useful to coders, such as the date, time code, subject number, trial number, and visual markers that appear at the points when the pictures and speech stimulus are presented to the child on each trial. Because coders are blind to trial type, they have no access to the visual stimuli seen by the child and they code eye movements in silence with no access to the verbal stimuli. Thus these markers are necessary for time-locking visual and auditory stimuli on each trial in relation to the child's eye movements.

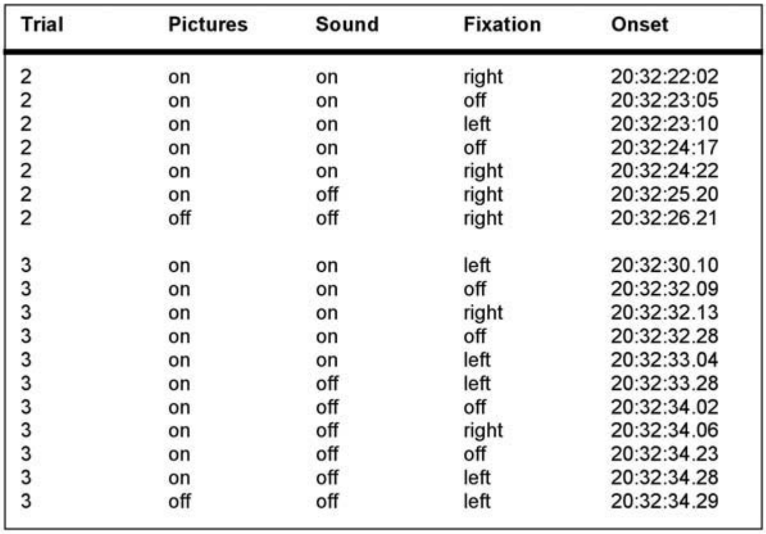
Pre-screening precedes coding and serves the primary purpose of flagging any non-useable trials on which the child's response was affected by factors unrelated to the auditory and visual stimuli. Two pre-screeners, blind to the hypothesis as well as to side of target presentation, independently watch each testing session in real time, with the sound on in order to identify any talking by the child or parent during the trial. Because the pre-screeners have access to the verbal

Figure 2. Screen shot of child's face during testing, showing information accessible to observers as they conduct frame-by-frame coding. The "Picture" and "Sound" codes appear on the screen as the images and speech sounds are presented to the child, enabling coders to link the record of eye movements to the onset and offset of visual and auditory stimuli across each trial.

stimuli, they are never the same individuals who code eye movements for that study. The task of the pre-screener is to mark in the *Eyecoder* record any trials that should be eliminated from the analysis of the experimental trials for any of the following reasons: the child was not looking at the pictures prior to sound onset; the parent or child was talking during the trial; the child was fussy or inattentive; the child changed position so the face was not visible, etc. Only after both pre-screeners have agreed on all useable and non-useable trials does eye movement coding begin.

2.8 Coding eye movements

Coding involves reducing the continuous record of the child's eye movements on each trial to a series of discrete events. Given that our goal is to measure the latency of the child's shift in gaze from the distracter to the named target picture in relation to critical points in the speech stream, as well as the duration of fixations to the target and distracter pictures on each trial, high-resolution coding is essential. Highly trained coders, unaware of trial type and target location, use *Eyecoder* to move frame-by-frame through the digital record of each test session. The coder starts by synchronizing the time code with the onset of the speech stimulus, marking the frame to which subsequent events are referenced. Figure 3 illustrates the time line of eye movements captured in the *Eyecoder* record, from the onset of the utterance to the offset of the picture on two trials. The coder's task is to



Trial	Pictures	Sound	Fixation	Onset
2	on	on	right	20:32:22:02
2	on	on	off	20:32:23:05
2	on	on	left	20:32:23:10
2	on	on	off	20:32:24:17
2	on	on	right	20:32:24:22
2	on	off	right	20:32:25.20
2	off	off	right	20:32:26.21
3	on	on	left	20:32:30.10
3	on	on	off	20:32:32.09
3	on	on	right	20:32:32.13
3	on	on	off	20:32:32.28
3	on	on	left	20:32:33.04
3	on	off	left	20:32:33.28
3	on	off	off	20:32:34.02
3	on	off	right	20:32:34.06
3	on	off	off	20:32:34.23
3	on	off	left	20:32:34.28
3	off	off	left	20:32:34.29

Figure 3. Sample coding record for two 4-s trials using *Eyecoder*. Each line indicates the time at which the coder judged that a change occurred, either in the stimuli, e.g., from *sound on* to *sound off*, and/or in the position of the child's fixation, e.g., from *right* to *off*.

advance the video record one frame at a time, indicating each time a change occurs in the child's fixation with one of four event codes: *left* (on left picture), *right* (on right picture), *off* (between the two pictures), and *away* (off-task, looking at parent). The coder also marks the frames when the visual cues for "sound" and "pictures" appear and disappear, indicating that the auditory and visual stimuli are *on* or *off*. Close attention is required to precisely code the first frame in which the infant initiates a shift from one object to the other and is currently not fixating either. Since our latency measures are based on such shifts, coders view each of these transition points several times by toggling back and forth between frames to be sure of their judgment.

2.9 Reliability coding

Because errors in identifying the onset and offset of visual fixations can have a substantial impact on the accuracy of reaction time (RT) analyses, coders undergo several weeks of training and practice, with close supervision. Each novice coder must complete a standard training set of four test sessions pre-selected to represent particular coding challenges in different age groups. They must achieve a reliability score of 96% or better on the training protocol before they are authorized to code actual data. After completing training, all coders continue to participate in regular reliability checks to assess the extent to which that all coders are "calibrated" according to lab standards. Coders are never informed that they are coding a session chosen for reliability purposes, to ensure that they are not being extra cautious because they know it is a reliability test.

2.9.1 *Lab-wide versus study-specific reliability assessments*

Lab-wide reliability checks ensure that no single coder deviates from the established coding norms of the entire lab. This involves selecting a single test session for all coders to judge independently, one in which the child shifts frequently or has significant head movement that makes eye gaze more difficult to judge; in other words, we select a test session in which the potential for disagreement among coders is high. Every coder is checked against every other coder to identify those individuals who may deviate from the lab standard by more than a single frame on any trial. If the reliability scores reveal a discrepancy, coders are re-trained to adhere to the norm, and any data they may have contributed since the last reliability check is reviewed and corrected as necessary. While lab-wide reliability tests examine agreement among all coders on a single testing session, *study-specific reliability* checks assess agreement among a subset of coders on several test sessions. This type of reliability assessment focuses on a specific data set to ensure consistency in coding within a particular study. Trials designated for study-specific reliability are chosen to provide the most rigorous reliability testing conditions for coders. We select 25% of the participants in a study for reliability checks, coding one quarter of those experimental trials on which two or more shifts occurred.

2.9.2 *Entire-trial versus shift-specific reliability scores*

The *Eyecoder* program enables comparison of the response line for each recorded event (i.e., picture and sound onset and offset as well as shifts in gaze), noting discrepancies that exceed one frame. This frame-by-frame comparison is summarized in two scores: an "entire-trial" percentage of agreement and a "shift-specific" percentage of agreement. The *entire-trial* agreement score is based on the percentage of frames on which two coders' judgments agree overall. Coders are considered to be in full agreement if they record an event on the same frame or differ from each other by no more than a single frame. This kind of overall agreement score is the traditional measure of reliability reported in many preferential-looking studies, and it may indeed be adequate when a global measure of looking preference is the major variable. However, an entire-trial agreement score does not provide a sufficiently rigorous measure of reliability when using more precise time course measures, given that the goal of such an analysis is to assess reliability at the same level of resolution as the measure of interest, in this case at the level of milliseconds. The problem is that the child is maintaining fixation on one or the other of the two pictures on a majority of the 90 consecutive frames on each 3-s trial. Since the likelihood of agreement is much higher during these sustained fixations than during periods of transition, it is possible for two coders to receive a very high entire-trial agreement score, yet still disagree about the exact frames on which shifts in gaze begin and end. This is problematic because judgments about the initiation and duration of shifts are crucial in measures of reaction time and accuracy. Thus the traditional entire-trial score alone is too lenient a measure of reliability, although in combination with a more stringent reliability measure, it can help identify discrepancies among coders. *Shift-specific reliability scores* focus only on sequences of frames where shifts occur. The shift-specific calculation is a percentage of frames from shift start to finish on which coders agree, differing by one frame at most. If entire-trial and shift-specific scores are below 95% and 90%, respectively, an independent, highly-experienced coder examines the discrepancies between coders and makes a final decision about the gaze patterns for the session under review. At that time, coders who may have shifted standards for judging eye gaze are re-trained as necessary.

2.9.3 *Data reduction using Datawiz*

When coding is completed for all the participants in a particular study, the *Eyecoder* data from each coded test session needs to be coordinated with information about the identity and side of the target object on each trial, and the onsets and offsets of critical words in the speech stimulus. Using the custom software *Datawiz*, the *Eyecoder* data from all participants in the same study are consolidated, at the same time integrating other relevant information about each child such as vocabulary scores into the record. The output from the *Datawiz* program is an Excel-formatted spreadsheet for each child, with a series of codes indicating the time course of gaze patterns to the target picture, distracter picture, or away/off of both pictures at every 33-ms interval for every trial. These codes are aligned relative to pre-determined critical words in the stimulus sentence in each condition.

Depending on where the child is looking at the onset of the critical word, each trial is classified as target-initial, distracter-initial, off (between pictures), or away. Figure 4 shows a schematic representation of several different types of response patterns over a series of hypothetical trials in which eye movements were time-locked to the onset of the target noun. The dashed vertical line indicates the average offset of a typical target noun, although in actual data this would vary from trial to trial, depending on the particular items used. The information shown Figure 4 is comparable to the *Datawiz* output, except that the child's gaze patterns within each trial are summarized here using solid or patterned bars rather than as codes at every 33-ms frame. As indicated by the dark solid bars, trials 1 and 2 are both *target-initial trials* on which the child correctly maintained fixation on

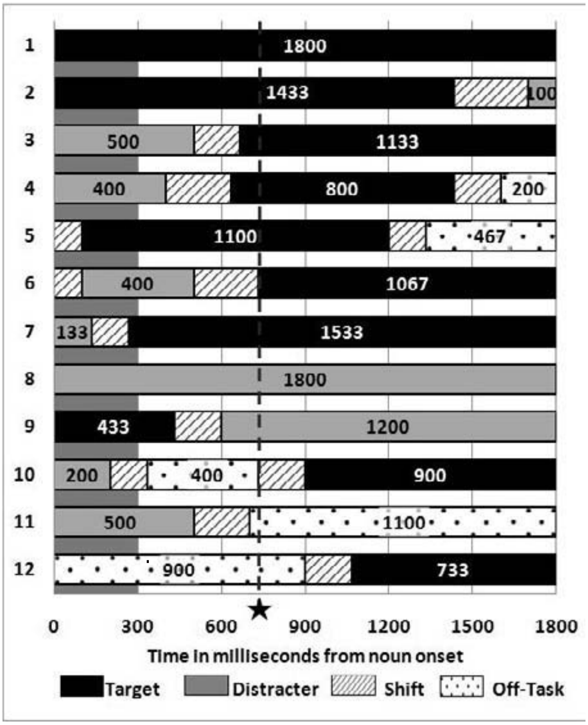


Figure 4. Schematic representation of different types of response patterns on hypothetical trials in which eye movements are time-locked to the onset of the target noun. The dashed vertical line indicates average target-noun offset. This sequence of trials is meant to illustrate different classes of possible response patterns and is *not* representative of the actual distribution of trial types observed in test sessions with particular children.

the correct picture for at least 1400 ms after the onset of the target noun. Trial 9 is also a target-initial trial; however, on this trial the child shifted away incorrectly to the distracter picture after hearing about 433 ms of the noun. Trials 3 and 4 are *distracter-initial trials* on which the child shifted quickly to the target picture, yielding reaction times of 500 ms and 400 ms respectively. On trial 7, the child also started on the distracter, but this time the shift was initiated only 133 ms after noun onset. Such very short shift latencies are not included in the calculation of mean reaction time, because they are likely to have been initiated prior to the point where the child had enough acoustic information from the noun to make an informed response and then to initiate an eye movement. On trial 8, the child again started on the distracter picture, maintaining fixation on the wrong picture for the duration of the trial, without ever shifting to the referent. On other trials in Figure 4, the child was either between pictures or off-task at the onset of the noun. Please note that the sequence of trial types shown here does *not* represent the distribution of response patterns for a typical experimental session with a child at an age for which the verbal and visual stimuli are appropriate. Trials 1–4 show the most common types of correct responses on distracter-initial and target-initial trials; trials 5–9 show patterns of correct and incorrect responding that are common, but occur less frequently, and trials 10–12 represent "away" trials that are relatively infrequent and would not be included in analyses of reaction time. Thus the goal in Figure 4 is not to show a representative sequence of trials for a particular child, but rather to provide examples of different gaze patterns that illustrate the variability possible in children's responses within the first 1800 ms following the onset of the noun.

2.9.4 *Data cleaning*

The output generated by *Datawiz* requires data "cleaning" steps before accuracy and RT can be calculated, to maximize the number of useable trials in the analyses. Recall that prior to the coding process, an experienced observer pre-screened each test session to exclude trials on which there was external interference or clear inattentiveness to the task. In the data cleaning process, the whole set of coded trials is further screened to evaluate off-task looking behavior *within* individual trials. The guidelines for accepting or rejecting individual trials are too detailed to describe here, but one example will suffice. On trial 3, the child started on the distracter picture and shifted to the target picture 500 ms from noun onset, in contrast to trial 11 where the child also started on the distracter and shifted away after 500 ms, but not to the target picture. In this case, the distracter-to-target shift in trial 3 would be included in the analysis of reaction time, while trial 11 would be excluded from the RT analysis because the shift in gaze was not directed toward either of the stimulus pictures.

3. Analyzing eye movement data from the looking-while-listening paradigm

This section examines different approaches to analyzing data on the time course of children's eye movements as they look at pictures and listen to speech that refers to one of the pictures. First we present plots of children's shifting gaze patterns in relation to particular words in the stimulus sentence as it unfolds over time. Next we describe how to derive discrete measures of children's speech processing efficiency from the time course information for use in statistical analyses, focusing on reaction time and accuracy. And finally, we address the questions of how stable such time course measures of processing efficiency are within individual children, and whether they are meaningfully related to other dimensions of early language competence.

3.1 Plotting graphs of the time course of children's eye movements in relation to speech

A useful first step in examining the data is to prepare an *onset-contingent (OC) plot*, which divides trials according to where the child is looking at the onset of the critical word in the stimulus sentence. An OC-plot tracks separately the time course of participants' responses for target-initial trials and distracter-initial trials as the stimulus sentence unfolds. At the beginning of a trial, the child has no way of knowing which object will be named, and so is equally likely to be looking at the target or the distracter picture at the onset of the target word. Thus the behavior that constitutes a correct response varies with the position of the child's eyes at the onset of the target word: on distracter-initial trials, the child should quickly *shift away* from the distracter to the named target picture; however, on target-initial trials, the correct response is *not* to shift but to stay put. The OC-plot in Figure 5 provides a graphic overview of these two different response patterns, using data from one condition in a cross-sectional study with children in three age groups: 18-, 24, and 36-month-olds (Zangl & Fernald 2007). For each participant, trials were grouped contingent on which picture the child was fixating at the onset of the target noun. Plotted on the y-axis, at each 33-ms interval from target word onset, is the mean proportion of trials on which children at that point are looking at a picture that is *different* from where they started at target-word onset. Because OC-plots capture children's tendency to shift away from the original starting point, they show both correct and incorrect responses: on distracter-initial trials, a shift away to the target picture is a correct response, while on target-initial trials, a shift away to the distracter picture is an incorrect response. Thus a child with perfect accuracy would shift quickly to the target picture on 100% of the distracter-initial trials, and would never shift away on target-initial trials.

The top three lines in Figure 5 track the mean proportion of *distracter-initial trials* on which children at each age have correctly shifted from the distracter and are now looking at the target at each 33-ms interval, plotted over participants from the onset of the target noun. The three lower lines track responses on *target-initial trials* for each age group, plotting the mean proportion of trials on which children have shifted away from the target at each time point and are now looking at the incorrect object. It is clear from the very different trajectories on distracter- and target-initial trials that the 36-month-olds were more likely than the younger children to respond correctly in *both* ways. On those trials when

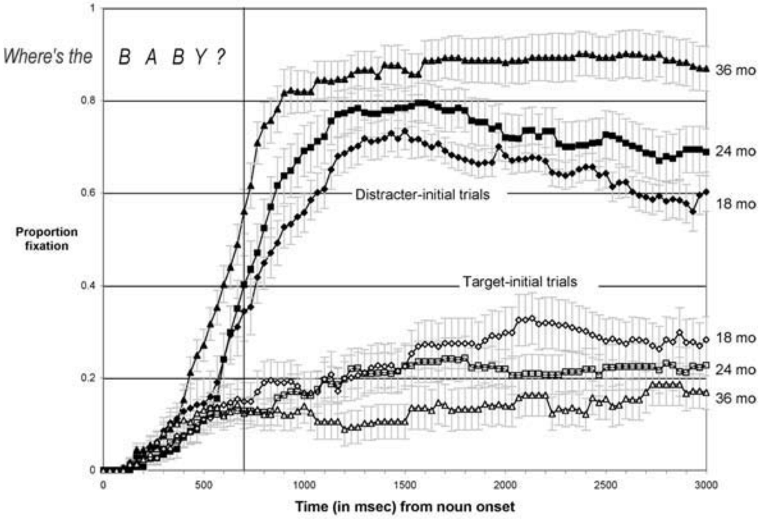


Figure 5. An *onset-contingent plot* (OC-plot) of distracter-initial and target-initial trials by children in three age groups, measured from target-word onset. At each 33-ms interval, data points show the mean proportion of trials either from the distracter to the target picture (on distracter-initial trials), or from the target to the distracter (on target-initial trials). On distracter-initial trials, a shift away to the target picture is a *correct* response, while on target-initial trials, a shift away to the distracter is an *incorrect* response. Error bars represent standard errors over participants (based on data from Zangl & Fernald 2007).

they started out on the distracter picture, the 36-month-olds began shifting to the correct referent halfway through the target noun. However, when they happened to start out on the target picture, they tended to maintain fixation on the correct referent rather than shifting away. In contrast, the 18-month-olds responded less efficiently on both counts: on distracter-initial trials, their gaze patterns suggested they were more likely to shift more slowly and less reliably to the target picture, and on target-initial trials, they were much more likely to false-alarm by shifting away from the correct picture. If children were *equally* likely to make distracter-to-target shifts and target-to-distracter shifts in response to a particular object name, this response pattern would suggest they were unable to identify the target word and/or to match it with the correct referent. But to the extent that children are quick and reliable to shift to the correct referent on distracter-initial trials, and *also* tend to maintain fixation on target-initial trials without shifting away, we can infer that they are able to interpret the spoken word efficiently and have associated the name with the right object. By providing information on both types of correct response, the OC-plot in Figure 5 offers a global view of developmental

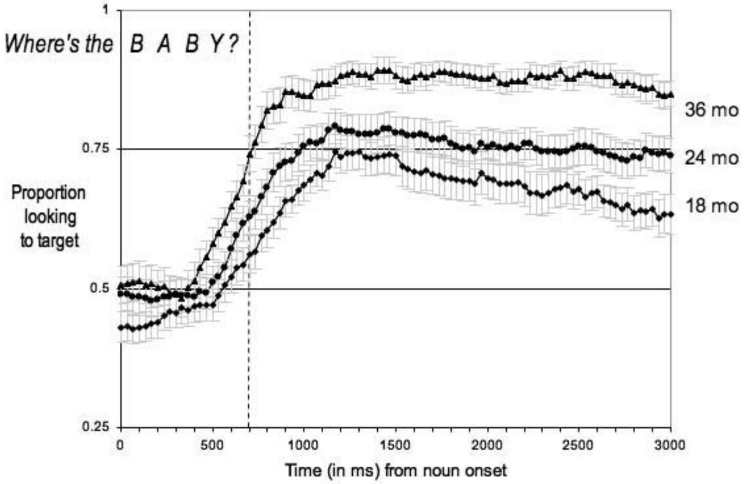


Figure 6. A *profile plot* showing the mean proportion of trials on which children in three age groups are looking at the target picture at each 33-ms interval as the stimulus sentence unfolds. Error bars represent standard errors over participants (based on data from Zangl & Fernald 2007).

changes in children's efficiency in identifying the appropriate referent as the target noun is heard.

A simpler way to graph the same data is to use a *profile plot*, as shown in Figure 6. This plot tracks the *mean proportion of looking to the target picture* at each time interval, measured from target-noun onset, averaged over participants for the same three groups of children as in the previous analysis. The overall results in Figure 6 are, of course, the same as those shown in Figure 5, with the 36-month-olds responding more reliably and reaching a higher asymptote than the younger children. Note that the curves are relatively flat at the beginning of the target noun, before children had accumulated enough phonetic information to identify the word. But visual inspection of the slopes of the three curves, and of the relative points at which they begin to rise, suggests that the 36-month-olds as a group responded more quickly to the target word than did the younger children. In contrast to OC-plots, profile plots *combine* responses from both target- and distracter-initial trials to show the overall mean proportion of trials on which the child is looking to the correct referent at each time point, regardless of whether the child was already looking at the target picture or had just shifted there from the distracter.

Figures 7 and 8 provide another example of how OC-plots and profile plots provide different vantage points on the same data. The data represented in these figures come from a pilot study of children's use of semantic cues to identify the referent of an upcoming noun (Fernald 2004). Here we asked whether 26-month-olds, like adults (Altmann & Kamide 1999), would respond more quickly to familiar nouns presented in sentence frames with a semantically related verb (e.g., *Eat the cookie*) than in frames with a semantically unrelated verb (e.g., *See the cookie*). Six object words were presented in both related and unrelated frames, carefully spliced to control for durations of carrier phrases and target words. Note that the graphs of children's responses in these two conditions are plotted from *sentence* onset rather than from target noun onset, because we predicted that the very first word in the sentence would have an influence on their speed of orienting to the target object. Indeed, on related-frame trials, 26-month-olds began orienting to the referent much sooner than on unrelated-frame trials; the OC-plot in Figure 7 shows that responses on target- and distracter-initial trials began to diverge as the children heard the verb, before the target word had been spoken. In contrast, on unrelated-frame trials children had to wait for the target noun, since no earlier cues were available to facilitate identification of the referent. In Figure 8, the same data are presented in a profile plot, showing that children looked more

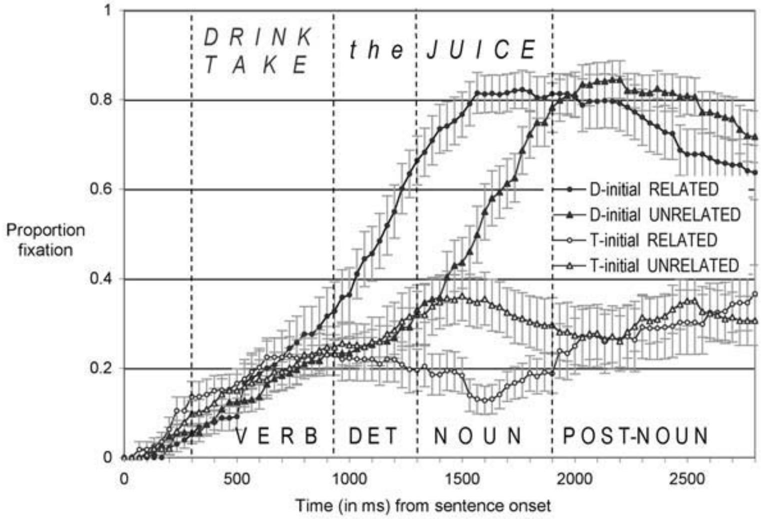


Figure 7. OC-plot showing at each time point the mean proportions trials on which 26-month-olds have shifted away from the picture they started on at sentence either correctly from the distracter to the target picture (on distracter-initial trials, top two curves), or incorrectly from the target to the distracter (on target-initial trials, lower two curves). Verbal stimuli were sentences with related and unrelated carrier frames preceding familiar target nouns. Dashed lines demarcate measurement windows for responses during the verb, determiner, and noun. Error bars represent standard errors over participants (adapted from Fernald 2004).

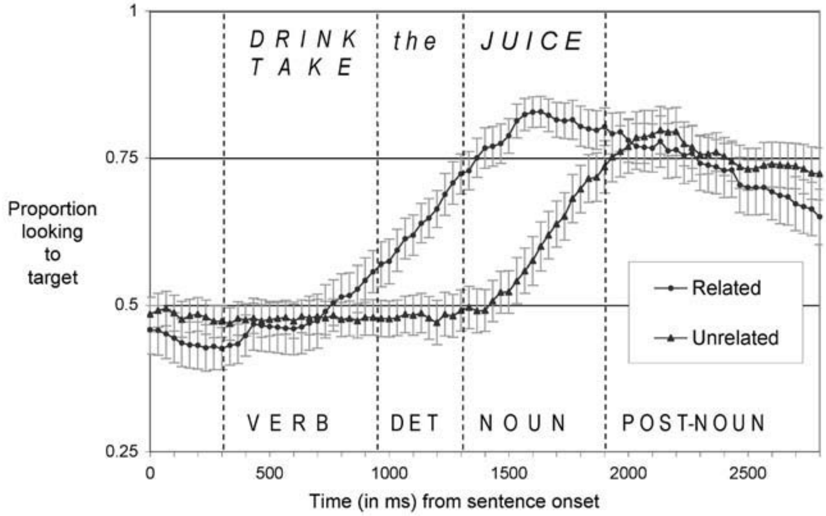


Figure 8. A profile plot showing the mean proportion of trials on which 26-month-olds fixate the target picture on related-frame and unrelated-frame trials. Dashed lines demarcate measurement windows for responses during the verb, determiner, and noun. Error bars represent standard errors over participants (adapted from Fernald 2004).

overall to the target picture on related- than on unrelated-frame trials. Unlike the data on age-related changes in word recognition shown in Figure 6, there was no difference in *asymptote* between the two conditions here, because 26-month-olds eventually converged on the target object to the same extent on both kinds of trials. What the profile plot in Figure 8 shows instead is a substantial difference in the *timing* of children's response between conditions. Like adults, 26-month-olds were able to take advantage of the semantic information in the verb to establish reference more quickly, but only on related-frame trials.

3.2 Deriving measures of processing efficiency from time course data

Although continuous plots of children's eye movement data provide a dynamic picture of the time course of their responses to particular words in the unfolding sentence, they do not directly represent the discrete measures that are most convenient for purposes of statistical comparison. The two dependent measures used most frequently in our research are *reaction time*, or latency to orient to the target word, and *accuracy*, based on looking time to the correct referent calculated over particular regions of the sentence.

3.2.1 *Reaction time*

A ubiquitous measure in psycholinguistic research with adults, reaction time (RT) has been used to explore how different linguistic and non-linguistic factors influence speed of lexical access and ease of sentence interpretation. Many such studies with adults have used experimental paradigms, such as lexical decision, that not only require participants to make a voluntary behavioral response, but also depend on metalinguistic judgments. Given these task demands, RT measures have not been widely used in research with infants, who have a very limited repertoire of voluntary behaviors that can serve as reliable response measures. However, moving the eyes to interesting stimuli is one behavior with which infants have extensive experience, and developmental researchers have found many ways to use infant gaze as a revealing experimental measure (Haith 1980). Although in the 1980's, a few studies of infants' sensory and perceptual abilities used fine-grained temporal measures to investigate the early development of visual scanning (Aslin 1981; Bronson 1982), researchers interested in cognitive development used more global measures of infant gaze patterns to reveal "looking preference" to one stimulus over another. A decade later, researchers studying visual cognition began to use looking behavior to measure response latency. Investigating how infants develop expectations when looking at a display of alternating visual stimuli, Haith, Wentworth, and Canfield (1993) established that the minimum latency for 3-month-olds to initiate a shift in fixation to a peripheral stimulus was around 200 ms. And when the infant had to disengage from one stimulus before initiating a shift to another, Hood and Atkinson (1993) found that responses were further delayed by 200 ms. These and other developmental studies of visual search (e.g., Canfield, Smith, Brezsnyak, & Snow 1997) were based on meticulous observations of infants' eye movements in a carefully controlled non-linguistic task, analyzed with millisecond-level precision. Coming from a completely different research tradition, these studies using visual RT measures with young infants converged with new research using eye movements to study spoken language understanding by adults (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy 1995), laying the foundation for the use of shift latency as a measure of processing speed in infant comprehension.

In the looking-while-listening procedure, RT is assessed on distracter-initial trials by calculating the latency of the infant's first shift away from the distracter toward the target picture, measured from a critical point in the stimulus sentence. Thus, the RT measure captures the point of departure from the initially-fixated distracter picture, not the point of arrival at the target picture, because we are interested in the decision to shift rather than the shift itself. Figure 9 shows the results of the first published study using RT measures of spoken word recognition by infants at 15, 18, and 24 months of age (Fernald et al. 1998). These results

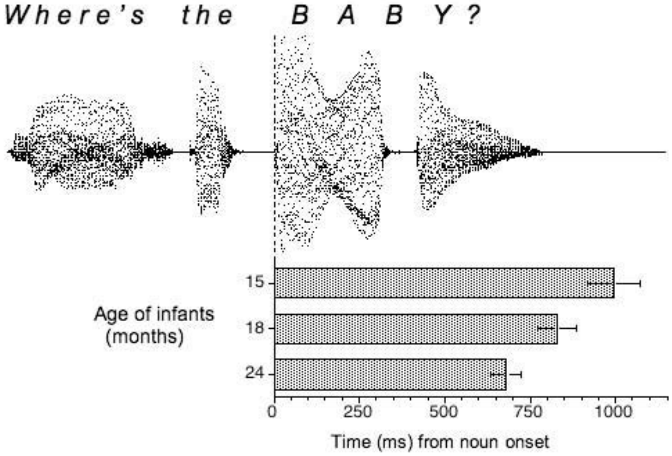


Figure 9. Reaction time analysis: mean latencies to initiate a shift in gaze from the distracter to the target picture, measured from the beginning of the spoken target word, for 15-, 18-, and 24-month old infants. This analysis included only shifts occurring on distracter-initial trials, i.e., those trials on which the infant was initially looking at the incorrect picture and then shifted to the correct picture within 1800 ms of target-word onset (adapted from Fernald et al. 1998).

revealed that over the second year of life, during the same period when most infants show a "vocabulary spurt" in speech production, they also make dramatic gains in receptive language competence by increasing the speed with which they can identify familiar words and match them with the appropriate referent. These cross-sectional findings have now been replicated in a longitudinal study with English-learning infants (Fernald et al. 2006), and we have also found comparable results with infants from Latino families living in the US, learning Spanish as their first language (Hurtado, Marchman, & Fernald 2007).

In the Fernald et al. (1998) study, all distracter-to-target shifts that occurred between the onset of the target word and the end of the trial were included in the RT analysis. However, in subsequent experiments we have been more selective, excluding very short response latencies that presumably reflect eye movements programmed before the start of the target word. Every distracter-to-target shift that is interpretable as a correct response has multiple components, each requiring processing time that can only be estimated. These components include the time required to accumulate enough phonetic information to identify the spoken word and then to assess its relevance to the currently fixated picture, and the time required to initiate an eye movement, if a shift is required. Based on the findings of Haith et al. (1993), it seemed reasonable to assume that infants in the looking-while-listening procedure need around 200 ms to program an eye movement, or perhaps longer given the need to disengage from an interesting distracter picture before shifting to the target (Hood & Atkinson 1993). Allowing time for 100 ms or so of phonetic information to accumulate, our first estimate was that it was reasonable to exclude from the RT analysis any shifts that occurred in the first 333 ms from the onset of the target word (Fernald et al. 2001); in more recent studies with older infants, we have used a cutoff of 300 ms (Zangl & Fernald 2007).

In addition to excluding very short latencies from the RT analysis, on the assumption that they were programmed before the target word was heard, it is also important to exclude very long latencies that are also unlikely to be in response to the target word. To develop guidelines for setting the boundary for such delayed shifts, we examined distributions of first shifts at different ages. The histograms in Figure 10 show the distributions of RTs on both distracter- and target-initial trials from a study of familiar word recognition by 18- and 21-month-olds (Fernald et al. 2001). Infants shifted to the target before the end of the trial on 88% of the distracter-initial trials; on 12% of these trials they never shifted at all. Note that the first shifts on six trials fell below the lower cutoff point of 333 ms, and thus were excluded from the RT analysis. In this data set we decided on 1800 ms as the upper boundary for shifts to be included in the RT analysis. This cutoff was chosen to eliminate responses influenced by the second repetition of the noun (since in this particular stimulus set each target word was spoken twice per trial), and it excluded outliers more than 2 SD greater than the mean of the distribution. For purposes of comparison, the distribution of response latencies on the 200 target-initial trials in this study is also shown in Figure 10B. Although infants shifted randomly to the distracter on some trials as the target word was spoken, more than half the infants did not shift at all, maintaining fixation on the correct picture.

The determination of appropriate cutoff points for excluding trials from an RT analysis is an important decision that may vary somewhat from study to study, depending on the experimental question and the age of the children in the study. The lower cutoff is relatively constrained; in eyetracking studies with participants of different ages, this cutoff has varied from 200–400 ms (Bailey & Plunkett 2002; Ballem & Plunkett 2005), with shorter intervals typically used with adults (e.g., Tanenhaus, Magnuson, Dahan, & Chambers 2000) and with children older than 24 months (Fernald et al. 2006; Zangl & Fernald 2007). Establishing the upper cutoff is not always as straightforward, although one reasonable approach is to identify outliers by examining the distribution of shifts for outliers, as shown in Figure 10.

One additional issue of concern in calculating response latencies is the problem of sparse data, always a risk in infant studies that have very few trials. It is important to keep in mind that RTs can *only* be calculated on those trials where

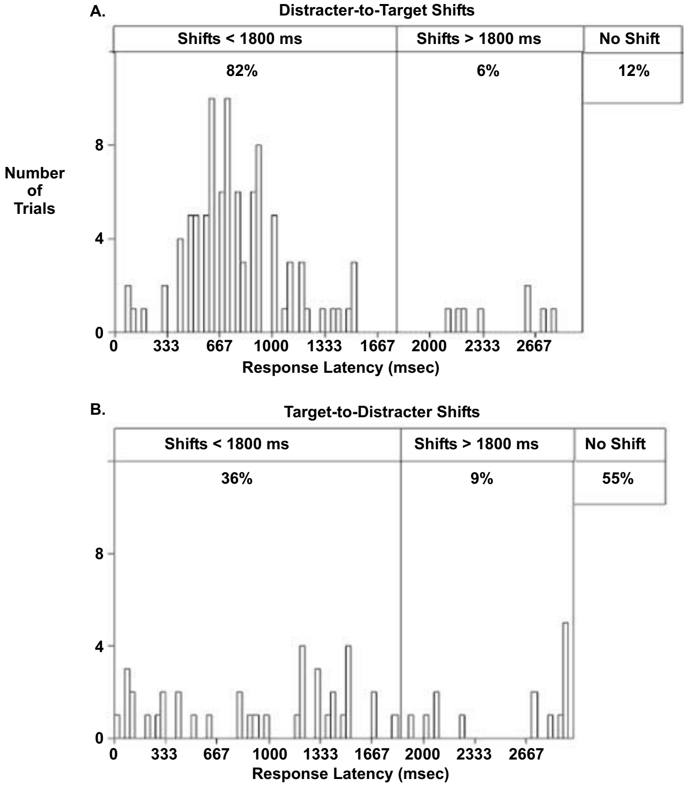


Figure 10. Distribution of RTs for first shifts from initially-fixated picture to alternative picture on 443 trials, following target word onset on (A) distracter-initial trials, and (B) target-initial trials. The proportion of trials on which no shift occurred within 3000 ms is noted on right (adapted from Fernald et al. 2001).

the child happens to start out on the distracter and then shifts to the target within the RT-window (e.g., 300–1800 ms from target-word onset), a subset of the data that often includes fewer than half the total number of trials. For this reason, in experimental designs with two or more within-subject conditions, it can easily happen that not all children contribute RT data and thus must be excluded from some analyses. For example, in a hypothetical experiment with 20 critical trials, 10 in each condition, each child will on average have eight useable trials, with two trials coded as "off" or "away" at target-word onset. Of the eight useable trials in each condition, some children may have six distracter-initial trials in each, all with shifts that occur within the appropriate RT-window. In this case the mean RT in each condition would be based on six trials – not an impressive number by the standards of studies with adults, but substantial for RT studies with infants. However, the more likely scenario is that children by chance will have only four distracter-initial trials in each condition, and some of these will be too fast or too slow to be included in the RT analysis. And of course some children by chance will have even fewer distracter-to-target shifts, and thus may end up with no RTs to contribute to the analysis at all. In earlier studies we have sometimes had to use mean RTs that were based on only two trials per condition, but this low criterion can result in very noisy data and works against finding positive results. To avoid this disappointing outcome, when designing an experiment in which latency measures are critical, it is important to make every effort to maximize the number of potential RT trials. This means including no more than two within-subjects factors, since each additional factor reduces the number of distracter-initial trials in each condition that will potentially yield RTs. Another approach is to double the overall number of trials by observing each participant in two separate sessions, scheduled a day or so apart.

3.2.2 *Accuracy*

Accuracy reflects how reliably children look at the correct referent, operationalized as the mean time spent looking at the target picture as a proportion of total time spent on either the target or the distracter picture, averaged over a particular region of interest. While the RT analysis is based only on distracter-initial trials, accuracy includes both target and distracter-initial trials, assessing looking time to the referent regardless of whether the child started out on the target picture or had to shift to the target from the distracter. Depending on the experimental question, accuracy may be measured across a single broad time window, or over multiple smaller time windows. For example, when assessing developmental changes in infants' accuracy in recognizing familiar words in simple sentence frames (*Where's the doggy?*), accuracy was calculated as the mean proportion of looking to the target over the broad window extending from 300 to 1800 ms from the onset of the target-word. This would be equivalent to the area under the curves over that window in the profile plot shown in Figure 6.

However, a multiple-window analysis was more appropriate for the study described earlier on children's use of semantic information from the verb to identify the referent (Fernald 2004). As shown in the profile plot in Figure 8, we defined four regions of interest: 1: verb, 2: determiner, 3: noun, 4: post-noun. Note that the measurement windows corresponding to these four regions incorporate the estimated 300 ms assumed to be necessary for processing the initial speech segments and mobilizing an eye movement; thus the verb window begins 300 ms into the verb and extends 300 ms beyond the offset of this word. Our prediction was that the accuracy curves on related-frame and unrelated-frame trials would begin to diverge at the end of the verb window, with significant differences emerging within the next few hundred milliseconds as children made use of informa-

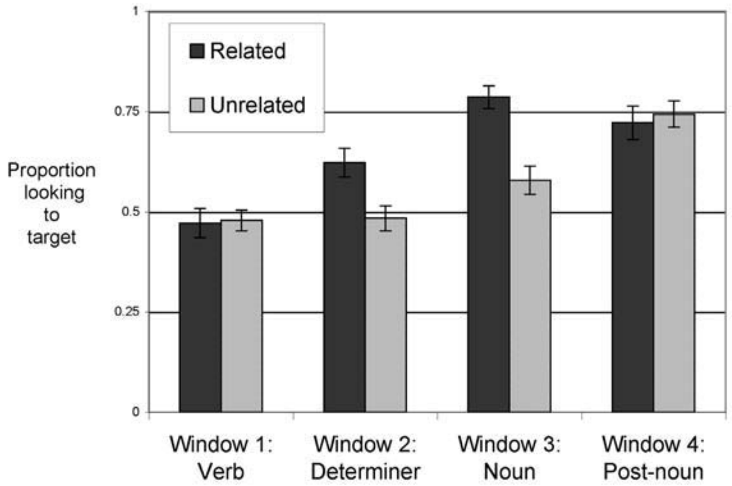


Figure 11. Accuracy analysis: Mean proportion looking to target averaged over participants in four critical windows: *Verb*, *Determiner*, *Noun*, *Post-noun*. During Window 2 and Window 3, 26-month-olds fixated the correct picture significantly more on related-frame trials than on unrelated-frame trials (adapted from Fernald 2004).

tion in the verb to find the target picture, even before they heard the noun. The accuracy analysis of these data is presented in Figure 11, showing the mean proportions of looking time to the target picture calculated over each of the four time windows. As predicted, the difference in accuracy between related and unrelated trials was significant in both the determiner and the noun windows, indicating that children could identify the correct referent using semantic information from the related verb. On related-frame trials, children were already fixating the appropriate referent 75% of the time, on average, by the beginning of the target noun; on unrelated-frame trials, in contrast, the mean proportion of looking time to the correct picture reached 75% only after the end the noun.

3.2.3 *Stability and predictive validity of on-line processing measures*

In the studies described here so far, on-line processing measures from the looking-while-listening procedure have been used to make between-group comparisons, tracking age-related differences in reaction time and accuracy in a cross-sectional design (Fernald et al. 1998), or examining condition differences in children of the same age (Fernald & Hurtado 2006; Thorpe & Fernald 2006) or between age groups (Lew-Williams & Fernald 2007). Using such group designs, we have shown that over the second and third year of life, children become faster and more reliable in recognizing familiar words in simple sentence frames, and that their ability to handle processing challenges such as morphosyntactic anomalies (Zangl &

Fernald 2007) and more complex sentence structure (Thorpe & Fernald under review) also improves dramatically over this period. These investigations of on-line processing efficiency by very young language learners provide new insights into the early development of receptive language competence, complementing results from studies of lexical and grammatical growth that are based on more traditional measures of speech production over the first three years.

But characterizing typical patterns of language growth over time is just one perspective in developmental research; another central goal is to characterize *variation* among children. For example, young children vary widely in the size of their productive vocabulary, and one 15-month-old may produce more than 50 words while another has not yet started to speak at all, with considerable variability apparent in grammatical as well as lexical growth over the early years (Bates, Dale, & Thal 1994). Research using on-line processing measures of language understanding can address important questions about differences among children within an age group, as well as between-group differences. Is speed of processing at any given age a *stable* measure for an individual child? That is, are children who respond more quickly on average in identifying familiar words at 18 months, relative to the mean RT for children at that age, the same children who respond relatively more quickly at later ages? And how do individual differences in efficiency of spoken language processing relate to individual differences in language growth, as assessed by standard measures of lexical and grammatical knowledge? In particular, does processing efficiency in infancy predict language and cognitive outcomes at later ages?

To begin to address these questions, we conducted a longitudinal study of 59 English-learning infants, testing them in the looking-while-listening procedure at 15, 18, 21, and 25 months of age (Fernald et al. 2006). Children's speed and accuracy in spoken word recognition increased significantly over this period, consistent with earlier cross-sectional research. To explore the relation of on-line measures of speech processing skill to more traditional measures of linguistic development, parental reports of vocabulary and grammatical usage were gathered at five time points across the second year, along with a standardized test of lexical knowledge at 25 months. Speed and accuracy in speech processing at 25 months were robustly related to lexical and grammatical development across a range of measures from 12 to 25 months. Analyses of growth curves revealed that children who were relatively faster and more accurate in spoken word recognition at 25 months were also those who had experienced faster and more accelerated vocabulary growth across the second year.

These findings led to the obvious next question: to what extent do individual differences in processing efficiency in infancy predict *later* language and cognitive outcomes? In a recent follow-up study (Marchman & Fernald under review), 30 of the children from the Fernald et al. (2006) longitudinal study were tested at the age of 8 years on the *Kaufman Assessment Battery for Children, Second Edition* (KABC-II) and the *Clinical Evaluation of Language Fundamentals, Fourth Edition* (CELF-4) standardized assessments of cognitive and language skills. Multiple regression analyses were used to evaluate the long-term predictive validity of two measures in infancy – expressive vocabulary and mean RT at 25 months – in relation to school-age outcomes. In light of the presumed links between efficiency of spoken language comprehension and working memory in older children and adults, we also examined relations between processing speed in infancy and performance working memory. Mean RT in the looking-while-listening task at 25 months was significantly correlated with scores on CELF-4 (r = –.52 to –.55) and KABC-II (r = –.43 to –.70) at 8 years, taking vocabulary size into account, and relations were strongest to performance on the working memory subscale. Moreover, it was children's speed in identifying the target word on *challenging* trial types, requiring them to integrate semantic or morphosyntactic cues, that accounted for the most variance; mean RT to familiar words in simpler frames did not add predictive power. This prospective longitudinal study is the first to reveal the long-term predictive validity of on-line measures of processing efficiency by very young language learners, showing that individual differences in the efficiency of spoken language interpretation at the age of two years predict children's success in cognitive and language tasks in later childhood.

4. Conclusions

It is fascinating and revealing to watch infants look at the world as they listen to speech in a carefully controlled experimental context. Their gaze patterns provide a window on their referential decisions, as they seek meaning in the words they are hearing in relation to the objects they are looking at, all within fractions of a second. A great deal of research on early language development aims to characterize what words children "know" at a particular age, as if words were "acquired" in an all-or-none fashion; however, the methods and results we have described here take a different perspective, focusing on the gradual development of children's efficiency in using their emerging lexical knowledge to interpret spoken language. According to this view, if the rather static notion of "acquisition" is appropriate to lexical development at all, then learning to make sense of a spoken word is like acquiring a skill rather than acquiring a thing, with an emphasis on gradual mastery rather than on possession. Infants may respond to more and more words over the second year, but they also learn to respond with increasing speed and efficiency to each of the words they are learning, and to recognize these words in more diverse and challenging contexts.

The looking-while-listening methodology described here uses fine-grained measures of the time course of children's gaze patterns in response to speech to explore the early development of language understanding. On the one hand, this procedure is technically "simple" in terms of stimulus presentation, similar to "preferential-looking" procedures in that it is low in task demands and does not require automated eyetracking technology. However, the looking-while-listening methodology differs critically from preferential-looking procedures in terms of the quantitative methods used for data reduction and analysis, yielding high-resolution real-time measures of speech processing rather than relying on summary measures of looking preference. Because gaze patterns are time-locked to the speech signal and coded frame-by-frame in this on-line paradigm, each 5-min experiment yields data from thousands of frames about the child's dynamic response to the unfolding sentence. As described in this chapter, the meticulous procedures involved in the collection, reduction, and multiple levels of analysis of such detailed data are certainly not simple. But they are well worth the effort, revealing a dynamic and nuanced picture of young children's developing skill in finding meaning in speech.

Acknowledgements

Thanks to Jonathan Berger, Nereyda Hurtado, Gerald McRoberts, Amy Perfors, John Pinto, Dan Swingley, and many other students and co-workers who have all contributed to the gradual refinement of the looking-while-listening paradigm described in this chapter. We are grateful to the National Institute of Child Health and Human Development (HD 42235) for their generous support of our research. Correspondence should be addressed to: Anne Fernald, Department of Psychology, Stanford University, Stanford, CA 94305.

References

Altmann, G. T. M. & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition, 73*(3), 247–264.

Aslin, R. N. (1981). Development of smooth pursuit in human infants. In D. F. Fisher, R. A. Monty, & J. W. Senders (Eds.), *Eye movements: Cognition and visual perception*. Hillsdale NJ: Lawrence Erlbaum Associates.

Bailey, T. M. & Plunkett, K. (2002). Phonological specificity in early words. *Cognitive Development, 17,* 1265–1282.

Ballem, K. D. & Plunkett, K. (2005). Phonological specificity in children at 1;2. *Journal of Child Language, 32*, 159–173.

Bates, E., Dale, P., & Thal, D. (1994). Individual differences and their implications for theories of language development. In P. Fletcher & B. MacWhinney (Eds.), *Handbook of child language.* Oxford: Blackwell.

Benedict, H. (1979). Early lexical development: Comprehension and production. *Journal of Child Language, 6,* 83–200.

Bloom, L. (1973). *One word at a time: The use of single-word utterances before syntax*. The Hague: Mouton.

Bronson, G. W. (1982). *The scanning patterns of human infants: Implications for visual learning.* Norwood NJ: Ablex.

Canfield, R. L., Smith, E. G., Brezsnyak, M. P., & Snow, K. L. (1997). Information processing through the first year of life: A longitudinal study using the visual expectation paradigm. *Monographs of the Society for Research in Child Development, 39* (No. 2, Serial No. 158).

Cutler, A. & Swinney, D. (1987). Prosody and the development of comprehension. *Journal of Child Language, 14*(1), 145–167.

Cutler, A. & Foss, D. (1977). On the role of sentence stress in sentence processing. *Language and Speech, 20*, 1–10.

Fantz, R. L. (1963). Pattern vision in newborn infants. *Science* 19 April 140 (no. 3564), 296–297.

Fenson, L., Marchman, V. A., Thal, D., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur-Bates communicative development inventories: User's guide and technical manual.* 2nd Edition. Baltimore MD: Brookes.

Fernald, A. (2002). Hearing, listening, and understanding: Auditory development in infancy. In G. Bremner & A. Fogel (Eds.), *Handbook of infant development* (pp. 35–70). Oxford: Blackwell.

Fernald, A. (2004). The search for the object begins at the verb. Presented at the 29th Annual Boston University Conference on Language Development, Boston, Nov. 4–7, 2004.

Fernald, A. & Frank, M. (in press). Finding the words: How young children develop skill in interpreting spoken language. In M. Spivey, M. Joanisse, & K. McRae (Eds), *Cambridge handbook of psycholinguistics.* Cambridge: CUP.

Fernald, A. & Hurtado, N. (2006). Names in frames: Infants interpret words in sentence frames faster than words in isolation. *Developmental Science, 9,* F33–F40.

Fernald, A., McRoberts, G., & Herrera, C. (1992). The role of prosodic features in early word recognition. 8th International Conference on Infant Studies. Miami, May 1992.

Fernald, A., McRoberts, G. W., & Swingley, D. (2001). Infants' developing competence in understanding and recognizing words in fluent speech. In J. Weissenborn & B. Höhle (Eds), *Approaches to bootstrapping in early language acquisition* (pp. 97–123). Amsterdam: John Benjamins.

Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology, 42*(1), 98–116.

Fernald, A., Pinto, J., Swingley, D., Weinberg, A., & McRoberts, G. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science, 9*, 72–75.

Fernald, A., Swingley, D., & Pinto, J. P. (2001). When half a word is enough: Infants can recognize spoken words using partial phonetic information. *Child Development, 72*(4), 1003–1016.

Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language, 14*, 23–45.

Grosjean, F. (1985). The recognition of words after their acoustic offset: Evidence and implications. *Perception and Psychophysics, 38*(4), 299–310.

Haith, M. M., Wentworth, N., & Canfield, R. (1993). The formation of expectations in early infancy. *Advances in Infancy Research, 8*, 251–297.

Haith, M. M. (1980). *Rules that babies look by: The organization of newborn visual activity.* Potomac MD: Lawrence Erlbaum Associates.

Halle, P. A. & de Boysson-Bardies, B. (1994). Emergence of an early receptive lexicon: Infants' recognition of words. *Infant Behavior and Development, 17,* 119–129.

Henderson, J. M. & Ferreira, F. (Eds.). (2004). *The interface of language, vision, and action: Eye movements and the visual world.* New York NY: Psychology Press.

Hollich, G., Hirsh-Pasek, K., & Golinkoff, R. M. (2000). Breaking the language barrier: An emergentist coalition model of word learning. *Monographs of the Society for Research in Child Development, 65*(3), Serial No. 123.

Hood, B. M. & Atkinson, J. (1993). Disengaging visual attention in the infant and adult. *Infant Behavior & Development, 16,* 405–422.

Hurtado, N., Marchman, V. A., & Fernald, A. (2007). Spoken word recognition by Latino children learning Spanish as their first language. *Journal of Child Language, 34*(2), 227–249.

Juszcyk, P. W. (1997). *The discovery of spoken language.* Cambridge MA: The MIT Press.

Katz, N., Baker, E., & McNamara, J. (1974). What's in a name? A study of how children learn common and proper names. *Child Development, 45*, 469–473.

Knoeferle, P. & Crocker, M. W. (2006). The coordinated interplay of scene, utterance, and world knowledge: Evidence from eye tracking. *Cognitive Science, 30*, 481–529.

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience, 5*(11), 831–841.

Lew-Williams, C. & Fernald, A. (2007). Young children learning Spanish made rapid use of the grammatical gender in spoken word recognition. *Psychological Science, 18,* 193–198.

Lewis, M. M. (1936). *Infant speech: A study of the beginnings of language.* London: Routledge & Kegan Paul.

Marchman, V. A. & Fernald, A. (2007). Speed of language processing and vocabulary knowledge in infancy predict school-age cognitive and language outcomes. Manuscript submitted for publication.

Markman, E. M. (1989). *Categorization and naming in children.* Cambridge MA: The MIT Press.

Marslen-Wilson, W. & Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 576–585.

Meints, K., Plunkett, K., & Harris, P. (2002). What is "on" and "under" for 15, 18- and 24-month-olds? Typicality effects in early comprehension of spatial prepositions. *British Journal of Developmental Psychology, 20*, 113–130.

Naigles, L. R. (2002). Form is easy, meaning is hard: Resolving a paradox in early child language. *Cognition, 86,* 157–199.

Portillo, A. L., Mika, S., & Fernald, A. (2007). Interpreting behaviors in "preferential looking" procedures: Do children seek a word/object match, or flee a mismatch, or both? Presented at the Biennial Meeting of the Society for Research in Child Development, Boston, March 29 – April 1 2007.

Reznick, J. S. (1990). Visual preference as a test of infant word comprehension. *Applied Psycholinguistics*, *11*, 145–166.

Saffran, J. (2002). Constraints on statistical language learning. *Journal of Memory and Language*, *47,* 172–196.

Shafer, G. & Plunkett, K. (1998). Rapid word learning by 15-month-olds under tightly controlled conditions. *Child Development, 69,* 309–320.

Shipley, E. F., Smith, C. S., & Gleitman, L. (1969). A study of the acquisition of language: Free responses to commands. *Language, 45,* 322–342.

Snedeker, J. & Trueswell, J. C. (2004). The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology, 49*(3), 238–299.

Spelke, E. S. (1976). Infants' intermodal perception of events. *Cognitive Psychology, 8*, 553–560.

Swingley, D., Pinto, J., & Fernald, A. (1999). Continuous processing in word recognition at 24 months. *Cognition, 71*(2), 73–108.

Tanenhaus, M. K., Magnuson, J. S., Dahan, D., & Chambers, C. G. (2000). Eye movements & lexical access in spoken-language comprehension: Evaluating a linking hypothesis between fixations and linguistic processing. *Journal of Psycholinguistic Research, 29*(6), 557–580.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*, 1632–1634.

Thomas, D. G., Campos, J. J., Shucard, D. W., Ramsay, D. S., & Shucard, J. (1981). Semantic comprehension in infancy: A signal detection analysis. *Child Development, 52,* 798–803.

Thorpe, K. & Fernald, A. (2006). Knowing what a novel word is not: Two-year-olds "listen through" ambiguous adjectives in fluent speech. *Cognition, 100,* 89–433.

Thorpe, K. & Fernald, A. (2007). Developing efficiency in on-line interpretation of adjective-noun phrases. Manuscript submitted for publication.

Tomasello, M. (2000). Do young children have syntactic competence? *Cognition, 74, 2*09–253.

Trueswell, J. C. & Tanenhaus, M. K. (Eds.). (2005). *Approaches to studying world-situated language use: Bridging the language-as-product and language-as-action traditions.* Cambridge MA: The MIT Press.

Trueswell, J. C., Sekerina, I., Hill, N., & Logrip, M. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition*, *73*, 89–134.

Walley, A. C. (1993). The role of vocabulary development in children's spoken word recognition and segmentation ability. *Developmental Review*, *13*, 286–350.

Werker, J. F. & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *7*, 49–63.

Zangl, R. & Fernald, A. (2007). Increasing flexibility in children's on-line processing of grammatical and nonce determiners in fluent speech. *Language Learning and Development, 3,* 199–231.

Zangl, R., Klarman, L., Thal, D., Fernald, A., & Bates, E. (2005). Dynamics of word comprehension in infancy: Developments in timing, accuracy, and acoustic degradation. *Journal of Cognition and Development, 6*, 179–208.

CHAPTER 5

What lurks beneath

Syntactic priming during language comprehension in preschoolers (and adults)

Jesse Snedeker and Malathi Thothathiri

How do young children represent the structure of an utterance? Do they employ abstract syntactic categories? Or are their representations more concrete and lexically limited? Our recent work brings together the world-situated eye-gaze paradigm and syntactic priming to explore these questions. We begin by reviewing theories of syntactic development and describing previous studies of syntactic priming during children's language production. Then we introduce our method for exploring priming during comprehension. Next we present a series of experiments on priming in adults, 4-year-olds and 3-year-olds. In each case the participants' interpretation is influenced by the structure of prior utterances, even in the absence of lexical overlap. We conclude that young children (and adults) employ abstract syntactic representations during on-line sentence comprehension.

1. Introduction

In the past ten years researchers have made extensive use of on-line methods to explore what children understand at different ages, how rapidly they understand it, and what kinds of information they use to reach this understanding. Our recent work brings together two existing methods (syntactic priming and the world-situated eye-gaze paradigm) to explore a different question: what is the nature and scope of young children's grammatical representations and how do they change over development? We begin by briefly reviewing alternate theories about the development of syntax and its mapping to semantics. Next we describe the phenomenon of syntactic priming during production, before introducing our method for exploring priming during spoken language comprehension. We then describe a series of experiments on syntactic priming from comprehension to comprehension in adults, 4-year-olds and 3-year-olds. In each case we find robust evidence that participants' interpretation of utterances is influenced by the structure of prior utterances even in the absence of lexical overlap. These results indicate that young children (and adults) employ abstract structural representations during on-line sentence comprehension. Finally, we discuss how this technique can be used to explore the nature of these abstract representations and their relation to individual lexical items.

1.1 The acquisition of argument structure: Early abstraction or item-based frames?

A perennial question in linguistics is how to characterize the relations between syntactic roles, semantic roles, and our knowledge of specific predicates (see e.g., Baker 1988; Goldberg 1995; Grimshaw 1990; Jackendoff 2002). Any theory of these relations must account for two facts. First, there are systematic linkages between meaning and syntactic structure that are robust across verbs and languages (Fillmore 1968; Baker 1988; Dowty 1991; Levin 1993). For example, agents of actions generally surface as subjects, and themes as direct objects. Second, although these linkages are systematic, they do not fully predict the syntactic position of an argument. Similar propositions can be expressed using different surface syntactic forms, depending on the verb, its morphological form and other factors such as discourse structure and the phonological weight of constituents (1)–(2). Since much of this variation depends upon the verb in the utterance, all viable theories make use of lexically-specific information (though they vary in whether it is syntactic or semantic and how it is used).

(1) a. The possible housing collapse frightened the young couple.
b. The young couple feared the possible housing collapse.

(2) a. Ariel sent a half-eaten pomegranate to Chris.
b. Ariel sent Chris a half-eaten pomegranate.

While every theory of language acquisition must acknowledge these two facts, theorists differ in which they see as primary. For one group of theorists, the robustness of the syntactic-semantic correspondences is seen as evidence that the linkages between meaning and structure are innate properties of universal grammar that play a role in language acquisition (Grimshaw 1981; Pinker 1984; Gleitman 1990). For example, Pinker's semantic bootstrapping hypothesis (1984) proposes that children come to the task of language acquisition with a set of thematic roles (e.g., agent and patient), a set of syntactic functions (e.g., subject and direct object), and some default rules for linking one to the other.

In contrast, usage-based theories place lexically-specific information at the center of acquisition (Tomasello 1992; Goldberg 1995). For example, Tomasello proposes that children initially analyze each predicate as an isolated grammatical island with open argument positions that can be freely filled with nominals. Gradually children begin to notice similarities in the semantic functions assigned to these fillers and their morphological marking or position relative to the verb. This observed overlap leads them to form broader semantic categories (such as agent and theme), broader syntactic categories (such as verb, subject and object), and generalizations about their relationship. This account differs from the semantic bootstrapping hypothesis in two ways. First, the semantic bootstrapping hypothesis proposes that children who are just beginning to learn language represent utterances in terms of broad semantic and syntactic categories that allow them to make generalizations from one verb to another. Second, the semantic bootstrapping hypothesis proposes innate default mappings between these semantic and syntactic primitives. These two features are partially independent. While innate mapping rules presuppose abstractions of roughly the same scope as the target grammar, the converse need not be true. Syntactic and semantic abstractions may guide children's early language acquisition even if the mappings between them must be learned.

In fact the notion that children's early grammars employ broad categories is shared by many theories which dispute the notion of innate or early-acquired thematic mapping rules. For example, Braine (1976), Bowerman (1973), and others have suggested that early child language is organized around conceptual categories (e.g., action, actor) which serve as an entry point into syntax. Goldberg (2006) emphasizes the learning of syntax-semantics mappings but nevertheless suggests that children have semantic generalizations (such as actor) and syntactic slots (like PP) from very early on. These theories all posit an early grammar with abstract categories that could support generalizations across verbs, thus they contrast with usage-based theories which claim early grammars lack such categories before 3- to 3.5 years of age.

Much of the recent research on these issues has addressed both questions simultaneously, searching for evidence that young children have thematic mapping rules that express relations between abstract syntactic and semantic categories. In contrast our work puts aside the question of whether young children have these linking rules and simply focuses on whether they have abstract categories. Specifically, we ask whether 3- and 4-year-old children show a form of structural priming that cannot be captured by a linguistic system that is limited solely to isolated, verb-specific representations. Before introducing the priming paradigm, we briefly discuss findings from other methods for exploring children's structural generalizations.

1.2 Novel-verb generalization as a window onto grammatical representations

Recent work on the nature of children's structural representations has focused on their ability to comprehend and produce sentences with novel verbs. An impressive number of production experiments has demonstrated that children under 3.5 primarily use new verbs in ways that mimic the input, failing to generalize argument structure alternations from one verb to another (see Tomasello 2000 for a review). For example, Tomasello & Brooks (1998) presented 2-year-old children with a novel verb in either a transitive or an intransitive construction (e.g., *The puppy is meeking the ball* or *The ball is meeking*) and then attempted to elicit the unmodeled construction. For example, intransitives were elicited by asking questions which placed the theme in subject position, ensuring that it would be given information in any response (*What did the ball do*?). Despite this discourse pressure, the children used the construction that had been modeled by the adult almost 90 percent of the time. The authors concluded that children's early constructions are verb-specific, and that abstract verb-general constructions develop gradually during the preschool years.

In contrast, several novel-verb comprehension studies have found evidence for abstract constructions in children between 20 and 36 months of age. These studies explore this generalization by testing whether young children can use broad semantic-syntactic mappings to interpret the thematic roles assigned by a novel verb. For example, Fisher and colleagues have found that children as young as 20 months systematically prefer to map transitive sentences to caused motion events (rather than self-generated motion events) but show no such preference for intransitive sentences (Fisher 2002a; Yuan, Fisher, & Snedeker 2007). Further evidence comes from children's comprehension of reversible transitives. By 21 months of age, children systematically interpret the subject of a transitive sentence with a novel verb as the agent of the action (Gertner, Fisher, & Eisengart 2006).

How can we reconcile the productivity present in these comprehension studies with the lack of generalization observed in the production studies? Most authors suggest that one set of findings reveals the child's workaday grammar, while the other reflects task-specific strategies or limitations. For example, Tomasello and colleagues have suggested that the preferential-looking studies may reflect emerging and incomplete representations that initially play little role in everyday comprehension and production (see, e.g., Savage, Lieven, Theakston, & Tomasello 2003). In contrast Fisher (2002b) has argued that low productivity during novel verb production cannot be taken as evidence for the absence of abstract representations. Whether a verb can appear in a particular argument structure alternation depends on a complex set of semantic constraints (Levin 1993). For example, the causative alternation, used by Tomasello and Brooks (1998), is restricted to verbs that encode an externally-caused manner of motion. Consequently, even a learner with abstract representations of argument structure might be unwilling to extend novel verbs to unattested constructions because she lacks full knowledge of the semantic constraints on the alternation or is uncertain about the meaning of the verb. While the children in these production studies clearly map the verbs to appropriate events, it is unclear how precisely their interpretation of the verb matches the one that the experimenters had in mind. Extracting the meaning of a novel verb from a visual scene is difficult even for adults (Gillette, Gleitman, Gleitman, & Lederer 1999; Snedeker & Gleitman 2004).

While the novel verb paradigms have been extremely informative, they have two limitations which have led researchers to seek out other methods. First, these paradigms necessarily explore children's structural representations by probing their knowledge of thematic linking rules. To generalize an argument structure alternation or interpret a sentence with a novel verb, a child must know how semantic roles are mapped onto syntactic positions. But as we noted earlier, while structural abstractions are necessary for adult-like linking rules, linking rules are not necessary for abstractions. Thus a paradigm which allows us to investigate abstractions without requiring knowledge of linking rules would be valuable.

Second, some have questioned the conclusions that can be drawn from novel-verb paradigms. For example, Ninio (2005) argues that children's ability to use or interpret novel verbs in unattested constructions does not necessarily demonstrate that they have linguistic representations which are abstract. Instead she suggests that children's grammatical knowledge is stored in a format that is lexically specific and concrete. However, under unusual circumstances, like those in the generalization studies, knowledge of one verb can be extended to another through a process of structural analogy.

The impact of this argument depends on our conception of analogy. Consider a child who hears "The bunny is gorping the duck" while watching two videos in which one actor pulls another by the feet. If she transfers structural knowledge from known verbs to the novel verb by virtue of the structure of the utterance or the fact that they are all *verbs*, then this process would posit precisely the kind of structural generalizations that the verb island hypothesis denies. However, it would also be possible to form such an analogy without invoking higher-level linguistic categories. Children could simply translate or substitute a novel verb with a known form that has the same apparent meaning (e.g., *gorp* means *pull*). This would allow them to apply item-specific knowledge without invoking larger generalizations. Thus knowing that knowledge is transferred from a known verb to a novel verb does not tell us about the nature of the representations that underlie this transfer or the knowledge that children draw upon in their comprehension and production of known words. Below, we discuss how syntactic priming paradigms can help shed light on these workaday structural representations.

2. Using syntactic priming to study abstraction in children's production

In an elegant and extensive series of experiments, Bock has demonstrated that, in adult speakers, prior use of a syntactic construction alters the probability that it will be used again (Bock 1986; Bock 1989; Bock & Loebell 1990; Bock, Loebell, & Morley 1992). For example, adults who have just produced a prepositional-object dative (PO e.g., *The girl handed a paintbrush to the man*) are relatively more likely to describe another picture using a second PO dative, while those who have just produced a double-object dative (DO e.g., *The girl handed the man a paintbrush*) are relatively more likely to describe the picture with a DO dative. This priming is *structural* in that it occurs even when the meanings of the prime sentences are controlled and the prime and target sentences have no content words in common. In addition to picture description, priming has been demonstrated during written sentence completion, spoken sentence completion, and dialogue with a confederate (Branigan, Pickering, Stewart, & McLean 2000; Branigan, Pickering, & Cleland 2000; Pickering & Branigan 1998). This priming is not due to lexical or prosodic similarity alone (Bock & Loebell 1990). The priming persists regardless of whether the speaker produces the prime or merely reads or hears it, demonstrating that these effects arise from the activation of structural representations or procedures that are common to both production and comprehension.

This technique has several advantages for studying the nature of children's structural representations. First, it allows us to explore how children use *known* verbs under controlled conditions. While novel-verb tasks are an effective way to control for prior verb-specific learning, interpreting their relevance for language processing may not be straightforward. As mentioned above, success at a novel-verb generalization task is compatible with reliance on lexically-specific representations for comprehension and production, supplemented by analogical problem-solving strategies. By looking for abstract structural priming with known verbs (for which children potentially could possess adequate lexically-specific representations), we can better evaluate the importance of abstract representations in children's everyday language use.

Second, this technique allows us to compare the relative strength of abstract and lexically-specific priming. We can assess this by comparing the priming effects under two conditions: (1) when there is no lexical overlap between the prime and target (e.g., prime: *The boy threw his dog a ball*, target: *The teacher gave the student new books*); and (2) when the prime and target sentences share the same verb (e.g., prime: *The boy gave his dog a ball*, target: *The teacher gave the student new books*). In adults priming is stronger when there is lexical overlap (Pickering & Branigan 1998 but see Konopka & Bock 2005). This suggests either that adults are using lexically-specific representations in addition to abstract ones, or that lexical items are connected to abstract representations by links which can themselves be primed (see Section 7.2). By comparing the strength of structural priming between verbs and structural priming within verbs, we can elucidate the relation between lexical and abstract processes in children and we can explore how both processes change over development. This may help reconcile the competing findings from the novel-verb production and comprehension studies.

Structural priming has only recently been used to study the nature of young children's representations. We are aware of four published studies that explore production priming in children. One examined the production of noun phrases in 3- and 4-year-olds using a confederate dialog paradigm (Branigan, McLean, & Jones 2005). The children showed strong abstract priming which was further enhanced when the prime and target shared the same head noun. While this study suggests developmental continuity in priming, it does not address the questions that motivate the current work. The controversy in language acquisition centers on the status of verbs in the grammars of young children. The psychological reality of nouns is uncontentious (for example, the verb island hypothesis posits a category of nouns that fill in the slots of lexically-specific verbal frames).[1]

The remaining three studies used a picture description paradigm. In this paradigm, participants are shown pictures of simple scenes. During the prime trials the scene is described for the participant, who is typically asked to repeat this description. The prime trial is immediately followed by a test trial during which the participant is simply shown a picture and asked to describe it. Critically, both the prime and test pictures depict events which could be described using two syntactically distinct forms (e.g., DO and PO datives or the passive and active forms of the transitive).

---

[1] The confederate dialog task used by Branigan and colleagues (2005) may tap different mechanisms than the picture description tasks described below. In adults, priming effects in dialog paradigms appear to reflect higher-level processes facilitating coordination during dialog, in addition to lower-level structural priming (Branigan et al. 2000; Pickering & Garrod 2004). The task used in the Branigan child study may have promoted direct comparison of the prime and target utterances, further encouraging parallelism. On every trial the experimenter and child each put down a card with a colored object on it, the experimenter described her card (e.g., *the red cat* or *the cat that is red*), the child described his, and then both participants raced to pick up the cards if they matched. The priming effects were larger than those observed in parallel studies with adults and prior studies with children (82% matches to the prime in the absence of lexical overlap).

The picture description studies with children differ from those with adults in several respects. First, studies with children typically employ a small number of test items. While this helps to ensure that the study is short enough to be completed before the child's attention wanders, it prevents the experimenters from ascertaining whether the observed effects are robust across items. Second, for the same reason, few if any filler items are used. Finally, the critical independent variables are often manipulated between subjects (rather than within subject) or blocked and presented in separate testing sessions. In adult studies, fillers and within subject manipulations are used to decrease the probability that participants will become aware of the critical manipulation and develop strategies specific to the experimental situation. Given the limited metalinguistic abilities of preschoolers (Gombert 1992), developmentalists are typically not concerned about this possibility. However, note that both of these changes could increase the amount of priming in the child studies. If all the primes are of the same type for a given participant and these trials are interrupted by few or no filler items, then priming from one trial may linger and summate with priming from the next.

The first study to explicitly explore priming in young children compared passive and active transitive constructions (Savage et al. 2003). Children were assigned to either a high or a low lexical overlap condition. In the high overlap condition, prime sentences used pronouns that could potentially be repeated in the target descriptions (e.g., *It got pushed by it*). In the low overlap condition, prime sentences used nouns that could not be repeated in the target descriptions (e.g., *The bricks got pushed by the digger*). 6-year-olds showed priming in both overlap conditions, but 3- and 4-year-olds showed priming in the high overlap condition only. Thus the authors concluded that while 6-year-olds have abstract representations, 3- and 4-year-olds primarily rely on lexically-specific representations involving pronouns and some grammatical morphemes. In contrast, Huttenlocher and colleagues found abstract structural priming in 4- and 5-year-olds for both transitive and dative constructions (Huttenlocher, Vasilyeva, & Shimpi 2004). While, these studies differed in several respects, two factors seem particularly relevant. The first is the number of times that the prime was repeated. The children in the Huttenlocher study heard each prime utterance just once, while those in the Savage study heard it four times. Repetition of a single utterance may primarily engage the mechanisms responsible for lexically-specific priming (e.g., verbal memory see Konopka & Bock 2005), resulting in little or no abstract priming. A subsequent study by Savage and colleagues provides some preliminary support for this possibility (Savage, Lieven, Theakston, & Tomasello 2006). They found that older 4-year-olds (mean age 4;11) showed weak priming effects when a single prime was used repeatedly, but showed robust priming when a variety of primes were employed. The second critical difference in the prior studies on abstract verbal priming is the exact age of the participants. The 4-year-olds in the Huttenlocher study were older than those in the Savage study (mean age of 4;8 as compared to 4;2). Thus the discrepancy could be explained if we assume that abstract production priming emerges sometime around four and a half. However, recent unpublished reports of production priming in younger children complicate this picture. Gamez, Shimpi, and Huttenlocher (2005) found no structural priming of datives in a picture description task with 3.5- to 4.5-year-olds, while Song and Fisher (2004) found robust structural priming in 3-year-old children using a sentence imitation task.

3. Studying priming during comprehension

Recent work in our lab explores syntactic priming during on-line comprehension. Since production tasks are often more difficult for children than comprehension tasks (Hirsh-Pasek & Golinkoff 1996), this may provide a more sensitive measure of children's linguistic knowledge and allow us to test younger children. These studies use a world-situated eye-gaze paradigm that taps on-line sentence processing (the visual world paradigm). We measure participants' eye movements while they listen to instructions and manipulate objects. Under such circumstances, eye movements to the objects are tightly linked to the unfolding utterances and are sensitive to lexical and structural processing in both adults (e.g., Allopenna, Magnuson, & Tanenhaus 1998; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy 1995) and children (Trueswell, Sekerina, Hill, & Logrip 1999; Snedeker & Trueswell 2004; Snedeker & Yuan in press; and Trueswell this volume). By using a technique with good temporal resolution, we can explore the locus of the priming effect and rule out alternate explanations that might apply to priming during production.

3.1 The "poor man's" eyetracker

Most researchers employing the visual world paradigm use head-mounted or table-mounted eyetrackers to measure fixation patterns (see Trueswell this volume). In our lab we use a method we call the "poor man's eyetracker", in which a hidden camera is used to videotape the participant's direction of gaze. The set-up is simple. The participant sits in front of an inclined podium with four shelves, one in each quadrant. A camera is placed beneath the podium with its lens aligned with a hole in the center of the display. The camera is focused on the participant's face and is used to record his or her eye movements, which are later coded using frame-by-frame viewing on a digital VCR.

There are several advantages to this method. First, we have found that more preschools, parents, and children are willing to participate in research that uses familiar and noninvasive technology. Second, the hidden camera technique is far less expensive than eyetracking, making it accessible to more researchers. All that is needed is a laptop computer with speakers, a small camera, a stage for displaying props, and some method for coding videotape frame-by-frame (e.g., an editing deck with a jog-shuttle knob or a computerized coding system). Since most of this equipment is already present in the typical language acquisition lab, the paradigm can be easily adopted by experimenters who are just beginning to explore on-line methods. Third, the equipment is light, compact, and very easy to set up, thus the paradigm does not require a dedicated testing room and can be readily transported to schools or distant populations. Finally the "poor man's" eyetracker avoids some of the technical limitations of other eyetracking methods. Because of their weight and fragility, head-mounted eyetrackers are not suitable for children under about four and a half. Table-mounted eyetrackers can be used with children of any age. However, most current models have difficulty tracking gaze if the child's head is in motion. Consequently, they are ill-suited for use during act out tasks and with children who have difficulty sitting still.

Although new to sentence processing, the "poor man's" eyetracker is simply a variant of the preferential-looking paradigms which are widely applied in developmental psychology (Fantz 1961; Fagan 1970; Spelke 1979). Intermodal preferential-looking studies typically show very high inter-coder reliability especially when frame-by-frame coding is employed (Hirsh-Pasek & Golinkoff 1996). Many of these intermodal studies have looked at children's comprehension of spoken language (Hirsh-Pasek & Golinkoff 1996). When frame-by-frame coding is synchronized with a speech stimulus, the paradigm is quite similar to the eyetracking paradigms used in sentence processing (see Fernald, Zangl, Portillo, & Marchman this volume). These techniques have proven to be sensitive enough to explore the resolution of pronouns in preschoolers (Song & Fisher 2002) and improvements in the speed of word identification between 15 and 24 months (Swingley, Pinto, & Fernald 1999).

In our lab we record and code eye movements using DVCAM equipment and tape stock. Unlike most formats DVCAM has audio-lock recording which ensures that the audio and video tracks remain synchronized. Coding is completed in two steps. The first coder listens to the tape with the audio on and notes the time of the onset and offset of the sentence. These time points are used to define the period during which the eye movements will be coded and to synchronize the eye movement data to the speech stream during analysis. A second coder views the tape with the audio off and notes the onset of each change in gaze and the direction of the subsequent fixation. The direction of a fixation is coded as being in one of the quadrants, at center, or away from the display. If the participant's eyes are closed or not visible, the frame is coded as missing and the data are excluded from the analysis. A subset of the tapes is independently coded by an additional observer and inter-coder reliability is generally high (> 90% in the studies presented below).

To validate this method, Snedeker and Trueswell (2004) performed a direct comparison of data collected with the hidden camera and data collected with a head-mounted eyetracker. They found that the two methods were quite comparable: they converged on the same fixation location for 93% of the video frames and produced similar amounts of lost data (2–3% of frames). Several aspects of our procedure may be critical to achieving this level of accuracy. First, the room is well lit and the camera is tightly focused on the participant's face, allowing the coders to see the iris and thus determine eye position. Second, participants are placed close to the display and their chair is positioned so that their gaze is centered at the location of the camera. This ensures that gazes to each of the four quadrants can typically be distinguished by the direction in which the eyes rotate and not merely by the extent to which they do so. The image of the participant's face on the hidden camera is monitored throughout the experiment to ensure that the participant remains properly positioned. Finally, to ensure that coders receive frequent feedback about the relation between eye position and gaze direction, we elicit a predictable sequence of gazes from the participants at the beginning of each trial by laying out the props in a consistent order (clockwise from the upper left) and drawing their attention to each one.

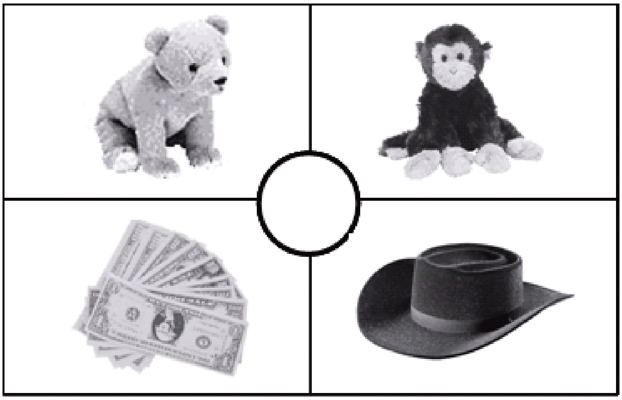


Figure 1. Example of a scene as viewed by the participant. Eye movements were recorded by a camera placed behind the hole in the center. DO sentence: *Bring the* ***monkey*** *the hat*; PO: *Bring the* ***money*** *to the bear*. (The ambiguous interval is in bold.)

3.2 Stimuli

In these studies we examine structural priming in sentences with dative verbs. Dative verbs, such as *give*, *bring*, or *send*, typically appear with three arguments: an agent, a recipient, and a theme. In English there are two ways in which these arguments can be expressed (see (3)). In the prepositional-object construction (PO, (3a)) the theme appears as the direct object while the recipient is expressed by the prepositional phrase marked by *to*. In the double-object construction (DO, (3b)) the recipient is the direct object while the theme is expressed as a second noun phrase and no preposition is used.

(3) a. The misanthrope left his entire fortune to Shamu.
b. The misanthrope left Shamu his entire fortune.

Datives are well-suited to our purposes for three reasons. First, they appear to be acquired quite early; children comprehend and produce both forms by age three (Campbell & Tomasello 2001; Gropen, Pinker, Hollander, Goldberg, & Wilson 1989). Second, the two dative constructions differ primarily in their syntactic structure and in the mappings between thematic roles and syntactic positions, and only slightly in meaning (if at all; see Baker 1997). Thus, priming using datives offers a reasonably clear case of structural priming independent of semantics. Finally, datives are commonly used in studies of production priming in adults, facilitating comparisons across experiments.

Each experimental block consisted of: (1) some filler sentences (which were not datives); (2) two prime sentences that were either DO or PO datives (e.g., DO: *Give the lion the ball*; PO: *Give the ball to the lion*); and (3) a final target sentence, which was also a DO or PO dative (e.g., DO: *Bring the monkey the hat*; PO: *Bring the money to the bear*). In all of our studies prime type and target type were fully crossed and manipulated between subjects, resulting in four possible conditions (DO prime-DO target; DO prime-PO target; PO prime-DO target; PO prime-PO target). Each participant in a given study was randomly assigned to one of these conditions.

Our goal was to determine whether DO and PO datives would prime the interpretation of subsequent utterances that used a different verb and had no common content words. To link this priming to eye movements we made use of a well-studied phenomenon in word recognition, the cohort effect (Marslen-Wilson & Welsh 1978). As a spoken word unfolds, listeners activate the lexical items that share phonological features with the portion of the word that they have heard. In the visual world paradigm, this process results in fixations to the referents of words that share features with the target word (Allopenna et al. 1998). These effects are particularly strong at the beginning of a word, when all of the phonological information is consistent with multiple words (the members of a given cohort). In our studies we used priming as a top-down constraint which might modulate the activation of different members of a phonological cohort.

On target trials, the set of toys that accompanied the utterance contained two items that were phonological matches to the *initial* part of the direct object noun (see Figure 1). One was animate and hence a potential recipient (e.g., a monkey) while the other was inanimate and hence a more likely theme (e.g., money). Thus the overlap in word onsets (e.g., *mon…*) created a lexical ambiguity which was tightly linked to a short-lived ambiguity in the argument structure of the verb. We expected that priming of the DO dative would lead the participants to interpret the first noun as a recipient, resulting in more looks to the animate match, while priming of the PO dative would lead participants to interpret it as a theme resulting in more looks to the inanimate match. The instructions were prerecorded by a speaker who used an enthusiastic tone and slow delivery.

3.3 Selecting a dependent variable

Dependent measures that are commonly used in eyetracking studies include first gaze duration, latency, and total fixation time. Young children's first looks may not be reliably guided by memory for a particular object in a particular location (Fernald, Thorpe, Hurtado, & Williams 2006). Therefore, we did not calculate first gaze duration or latency, and analyzed total fixation time only. We will refer to total fixation time simply as *looks*. In most of these studies our analyses focused on the interval during the target trials in which the identity of the direct object, and hence the argument structure of the verb, was temporarily ambiguous (e.g., *mon…*). Within this interval, we were interested in looks to the potential animate recipient (e.g., monkey) and the potential inanimate theme (e.g., money). We will refer to these two items as *animal* and *object* respectively. Specifically, we were interested in whether the type of prime sentence influenced how much the children looked to either the animal or the object. We explored three different dependent variables: (1) looks to the animal as a proportion of all looks; (2) looks to the object as proportion of all looks; and (3) the difference between the proportion of looks to the animal and the object.

Because eye movements are influenced by factors other than the ones manipulated (e.g., visual salience and name frequency; see Henderson & Ferreira 2004 for discussion), looks to one of the two items (animal or object) may be higher than looks to the other irrespective of the experimental condition. Thus, ceiling or floor effects might lead us to find significant effects for one of the measures but not the others. In our pilot studies we found that analyses of the *dispreferred* item appeared to be more sensitive to priming effects. For example, if participants preferred to look at the animal irrespective of the experimental condition, then we would find reliable effects of priming in the analysis of looks to the object but not in the analysis of looks to the animal. Therefore, our primary measure in the studies that follow will be looks to the dispreferred item.

4. Verifying syntactic priming in adult comprehension

There was, unfortunately, one road block to using comprehension priming to explore the development of structural generalizations. The existence of this form of priming is controversial in adults, raising the possibility that priming paradigms are insensitive to structural representations in comprehension, or that abstract syntax plays a weaker role in comprehension than production (Townsend & Bever 2001). Critically, while several prior studies have explored structural priming during comprehension, none of them provide unambiguous evidence of abstract syntactic priming for the kinds of sentences that we intended to explore (post-verbal ambiguities in argument structure). Two of the studies that found robust priming effects (Luka & Barsalou 2005; Noppeney & Price 2004) used structures which were not semantically equivalent (e.g., relative-clause attachment ambiguities), leaving open the possibility that the priming effects were semantic rather than syntactic. In addition, these studies used measures with a coarse temporal grain, creating uncertainty about whether the effects were due to initial structural analysis or later reanalysis. Two recent eyetracking studies have searched for evidence of priming during on-line comprehension. Scheepers and Crocker (2004) studied the processing of German transitive sentences with case marking. They found that the on-line interpretation of ambiguously marked preverbal arguments was influenced by prior unambiguously marked prime sentences. In contrast, using the English dative alternation, Arai, Van Gompel, and Scheepers (2007) looked for priming of the interpretation of post-verbal arguments. They found priming when prime and target sentences contained the same verb, but not when they contained different verbs (see also Branigan, Pickering, & McLean 2005).

Thus, while there is robust evidence for verb-specific priming during comprehension the scope and status of abstract priming is uncertain. Our first experiment revisited the role of abstract structural information during adult language comprehension. We tested 28 undergraduates in a between-verb priming task. In this study, we used a priming paradigm in which each sentence was acted out (act-out priming task, hereafter). Each participant heard four blocks of instructions. In each block, the first two sentences were fillers (non-datives), the next two were DO or PO prime dative sentences, and the last was a target DO or PO dative sentence containing a temporary ambiguity (e.g., money/monkey). For every sentence, new toys were put out, a sound file with an instruction was played, and the participant carried out the command. Thus from the participant's perspective there was no obvious difference between the filler, prime, or target trials. *Hand*, *pass*, *feed*, and *send* were each used in two prime sentences in two different blocks. *Throw* and *show* were each used in two target sentences.

Unsurprisingly, these adults performed the right action on all target trials, indicating that they were able to interpret the utterance irrespective of the prime type. However, their eye movements during the period of ambiguity were clearly affected by the prime. Since there was a weak preference for looks to the animal, our primary analysis focused on looks to the object (Figure 2, left panel). Participants who had heard PO primes were more likely to look at the object (the potential theme) than those who had heard DO primes. There were no reliable effects of prime type on looks to the animal (the preferred item). The effect of prime type persisted in the analyses of the difference scores. While participants who had heard DO primes clearly preferred the animal (the potential recipient), those who had heard the PO primes had a weak preference for the object.

Because the prime and target sentences in this experiment used different verbs and nouns, these results suggest that abstract representations are used during on-line language comprehension. Thus they contrast strongly with the results of Arai, Van Gompel, and Schepeers (2007) who found no between-verb comprehension priming for datives. That study used a passive viewing task, while our task required participants to plan and execute an action. Thus we considered the possibility that the locus of our priming effect was in the mapping from the utterance to the action plan. We explored this by testing whether abstract priming

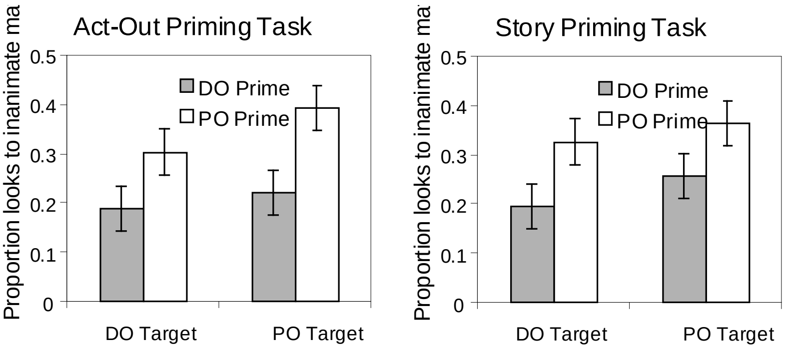


Figure 2. Structural priming effects during comprehension in adults: Proportion of looks to the inanimate match (consistent with PO priming) in the act-out priming task (left panel) and the story priming task (right panel).

persists when prime sentences are not enacted or mapped onto a visual array. If priming persists under these conditions we can rule out the possibility that actions are critical to producing priming during comprehension.

Participants passively listened to a story that contained the prime sentences and then acted out a target sentence using a set of props (story priming task, hereafter). Specifically, the participants were told that they were going to listen to two voices (Bob and Susan) which would be presented on a computer. Bob would tell stories, while Susan would give them instructions to act out. At the beginning of each trial, Bob's voice talked about events that happened in a children's store the day before. Susan's voice then interrupted with *It's my turn. Are you ready?* followed by the actual instruction. The first three trials were practice items that did not involve datives, while the subsequent trials alternated between critical trials and filler trials. On critical trials, Bob's last two sentences were DO or PO dative primes, while Susan's target instruction was also a DO or PO dative. There were a total of 6 critical trials with *read, teach, sing, show, sell,* and *feed* as the prime verbs and *bring, pass, throw, send, toss,* and *hand* as the target verbs. As Figure 2 illustrates, the effects of priming persisted in the story priming task, demonstrating that priming occurs even when the prime utterance is not mapped to an action or a visual display. Again there was a weak preference for looks to the animal and so our analyses focused on looks to the object. Participants who had heard PO primes were more likely to look at the object than those who had heard DO primes.

The two experiments reported here demonstrate comprehension-to-comprehension priming in adults when different verbs are used in prime and target sentences. These studies extend the findings of the prior comprehension studies in several ways. First, they demonstrate that priming occurs even when semantically equivalent dative sentences are used, thus minimizing the possibility that the effects are semantic rather than syntactic. Second, they show that priming unfolds soon after the onset of the first noun, which was on average less than 550 ms after verb onset. This suggests that priming influences initial syntactic analyses. In addition, these results complement Scheepers and Crocker (2004) by showing priming during the interpretation of post-verbal arguments. Abstract, non-verb-specific information appears to influence comprehension even after a specific verb has been encountered.

We attribute the divergence between our findings and those of Arai, Van Gompel, and Scheepers (2007) to two differences between the studies. First, Arai and colleagues used a single prime before each target trial while we used two. Previous evidence suggests that encountering multiple verbs in a structure leads to stronger structural priming (Pickering & Branigan 1998; Savage et al. 2006). Second, Arai and colleagues presented the target sentence immediately after the prime with no intervening verbal materials, resulting in a lag of approximately

1200 ms between prime and target trials. The lag in our studies was considerably longer, in terms of both intervening utterances and elapsed time (approximately five sentences and 30–60 seconds for the act-out task and two sentences and 4–5 seconds for the story priming task). A recent study by Konopka and Bock (2005) suggests that the distance between the prime and the target affects the relative magnitude of lexically-specific priming and abstract priming. They found lexically-specific priming only when the target immediately followed the prime. In contrast, abstract priming was numerically greater when a single sentence intervened between the target and prime (priming at lag 1 > priming at lag 0) and remained robust across as many as three intervening sentences. Konopka and Bock attribute lexically-specific priming to an explicit memory for the prime sentence which decays rapidly. Abstract priming, they argue, involves a form of implicit learning, parallel to the setting of connection weights in a neural network (Chang, Dell, & Bock 2006). Extending this proposal to comprehension priming generates the prediction that abstract priming would be greater in the present experiments while lexically-specific priming would be greater in the study by Arai and colleagues. Systematic investigation of the effect of the prime-to-target lag on comprehension priming will be required to validate this speculative account.

5. Syntactic priming in preschoolers

Armed with a paradigm that was sensitive to abstract structural priming in adults, we set out to discover whether the same paradigm could be applied to children. We focused our work on two age groups: young 4-year-olds and young 3-year-olds. Young 4-year-olds were of interest because they have failed to show abstract priming effects in two production priming studies (Savage et al. 2003; Gamez et al. 2005) despite showing fairly robust generalization in novel-verb production tasks (Tomasello 2000). This raises the possibility that novel-verb generalization paradigms may not reflect the structure that underlies everyday language use (Ninio 2005). Young 3-year-olds were of interest because they typically fail to generalize in novel-verb production studies (Tomasello 2000). Thus evidence of abstract priming in this age group would challenge the empirical basis of the verb island hypothesis.

Of course, one cannot simultaneously test the sensitivity of a method and the existence of the phenomenon that it is supposed to be sensitive to. Failure to find parallel effects of abstract structural priming in young children could indicate either that the task is not appropriate to explore priming in this age group, or that children fail to employ abstract representations in on-line comprehension. Thus we began by testing our task on an uncontroversial phenomenon: within-verb priming in 4-year-olds. Within-verb priming can be mediated by abstract structure or lexically-specific representations. Given the results of the prior production priming studies and the novel-verb generalization studies we would expect to find robust priming within verbs in any task that is sensitive to priming of the relevant representations.

5.1 Experiments with 4-year-olds

Twenty young 4-year-olds (M = 4;1) participated in an act-out priming task, similar to the one we conducted with adults. All prime and all target sentences used the verb *give*, which is the most frequent dative verb in the input to children and in speech between adults. The temporary ambiguity in each target sentence was created by using an animal as the recipient in the DO sentences and a compound noun beginning with the same word as the theme in the PO sentences (e.g., DO: *Give the bird the dog bone*; PO: *Give the birdhouse to the sheep*), resulting in a long ambiguous region (400 ms).

The children performed the right action on 89% of the target trials. Two-thirds of the errors were role reversals (e.g., giving the bird *to* the dog bone in response to *Give the bird the dog bone*). Most of these were in the mixed conditions, where the prime type did not match the target type, suggesting that children were sometimes led down the wrong path by the prime sentences. Across conditions, children looked at the animal more than the object, thus our analysis focused on looks to the object (Figure 3, left panel). Those children primed with PO sentences (where the first noun is the inanimate theme) looked more at the object than those primed with DO sentences (where the first noun is the animate recipient).

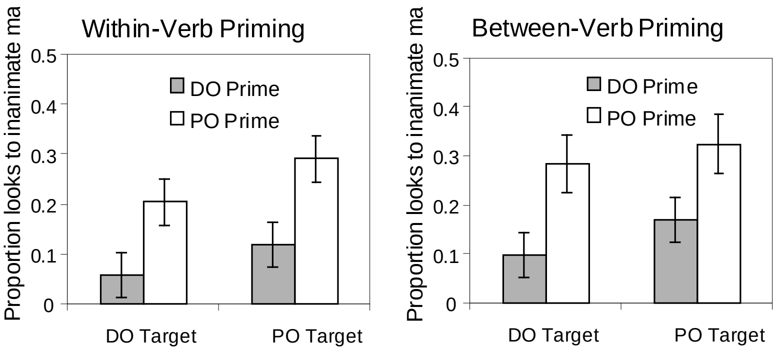


Figure 3. Structural priming effects in 4-year-olds: Proportion of looks to the inanimate match (consistent with PO priming) in the within-verb priming condition (left panel) and the between-verb priming condition (right panel).

Thus we found that 4-year-old children's interpretation of temporarily ambiguous dative sentences was rapidly influenced by the previous sentences that they had heard. Because we used the same verb in both prime and target sentences, this effect could reflect either verb-specific or abstract priming.

To explore whether 4-year-olds have structural representations that are broader than individual verbs, we conducted a parallel study of across-verb priming. The prime sentences used *show* and *bring,* while the targets were the same as those used in the within-verb priming study, and thus used the verb *give*. Thirty-eight young 4-year-olds participated (M = 4;0). The children performed the correct action on 90% of the target trials. Our analysis of looks to the dispreferred item (the object) revealed a significant effect of prime (Figure 3, right panel). As predicted, those primed with PO sentences looked more at the object than those primed with DO sentences. Thus we found that 4-year-old children's interpretation of temporarily ambiguous *give* sentences was influenced by the previous *show* or *bring* sentences that they had heard. This priming across verbs demonstrates that 4-year-olds have structural representations of dative utterances that are not bound to individual verbs.

There are however, two limitations to this finding. First, while these results clearly demonstrate between-verb priming for the target verb *give*, they cannot tell us whether this priming occurs across a wider range of dative verbs. The verb *give* is unique: it is the most frequent dative verb in the input and under some theories it has a privileged role in the acquisition of dative constructions (Goldberg 1995; Ninio 1999). Second, although the findings of the production priming tasks have been mixed, most theorists would agree that 4-year-olds demonstrate some degree of abstract structural generalization in novel-verb production tasks. Much of the debate about the nature of children's representations has centered on 3-year-olds. Thus it was imperative to extend these findings to younger children and to a wider variety of dative verbs.

5.2 Experiments with 3-year-olds

To validate our technique with a younger age group we began by exploring within-verb priming using the act-out priming task. Thirty young 3-year-olds (M = 3;1) participated. Each child heard one of two lists each containing four different dative verbs, but within a given block, the prime and the target verbs were the same. The verbs were *pass, send, throw,* and *bring* for group 1 and *hand, show, toss,* and *take* for group 2. The phonological ambiguities in target sentences did not depend upon compound nouns, because we were unsure whether 3-year-olds would know them. Instead we used animal/object name pairs that overlapped in their onsets (e.g., DO: *Show the* ***horse*** *the book*; PO: *Show the* ***horn*** *to the dog*).

Because the 3-year-olds were slower and more variable in their eye movements than the 4-year-olds, we averaged looking time over a longer time window in our analyses. While minimum saccade latencies can be as small as 133 ms for adults in a simple visual task (Matin, Shao, & Boff 1993), the latencies for young children in a task where all stimuli stay visible throughout the trial and there is phonological overlap amongst the visible items are likely to be higher. For example, Swingley, Pinto, and Fernald (1999) found mean latencies of 558 ms and 785 ms for adults and 24-month-olds respectively. In our experiments, average latencies to look at the first mentioned item on ten randomly selected, unambiguous prime trials were 983 ms for the 3-year-olds (SD = 292 ms) and 437 ms for the 4-year-olds (SD = 188 ms). Therefore, our analyses for 3-year-olds began 200 ms after the onset of the first noun but (conservatively) extended up to 2 seconds after noun onset. Because this window is likely to include looks that were programmed after the first noun was disambiguated, we might expect to see an effect of target type in addition to any prime effects.

Children performed the correct action on 79% of the target trials. Most errors (72%) were due to children not acting out DO sentences (picking up the toys but not carrying out the action). Across conditions, children showed a slight preference for the object over the animal. Our analysis of looks to the dispreferred item (the animal) found a significant effect of prime (Figure 4, left panel). As predicted, those primed with DO sentences looked more at the animal. Unsurprisingly, there was also a significant effect of target, reflecting the disambiguation of the direct object during this time window. In addition there was a marginal interaction between prime and target, suggesting that the priming effect was stronger for DO target sentences.

The error rate in 3-year-olds was considerably higher than in 4-year-olds. Because eye movements on error trials are hard to interpret, we performed a secondary analysis excluding those trials where there was an error in the action. The effect of prime type on looks to the dispreferred item (the animal) persisted in this analysis. Thus we found within-verb priming in 3-year-old children using eight different dative verbs. Those primed with DO sentences looked relatively more at the animal than those primed with PO sentences. These results demonstrate that priming is not restricted to frequent, prototypical dative verbs such as *give*.

Because this within-verb priming could arise from either verb-specific or more abstract representations, our next experiment examined between-verb priming in this population. Thirty-two young 3-year-olds (M = 3;1) participated. Each child was assigned to one of two stimulus lists. In one list *pass, send, throw,* and *bring* appeared as targets while *hand, show, toss,* and *take* appeared as primes.

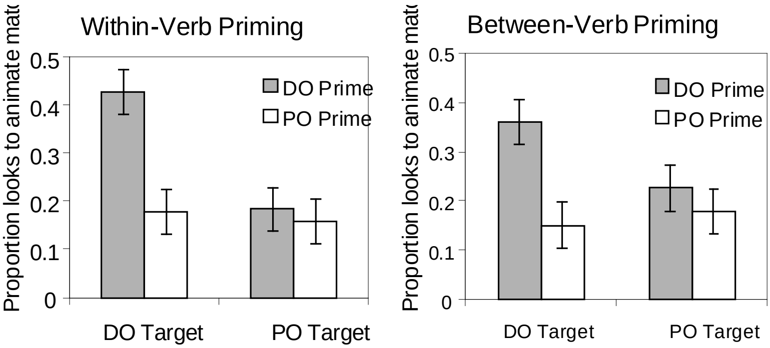


Figure 4. Structural priming effects in 3-year-olds: Proportion of looks to the animate match (consistent with DO priming) in the within-verb priming condition (left panel) and the between-verb priming condition (right panel).

In the second list, the role of the verbs was reversed. Thus the target and prime sentences were the same as those in the within-verb priming study but they were simply paired differently across participants. The two prime sentences in each block used two different verbs because varied primes have been found to lead to greater production priming in older children (Savage et al. 2006).

Children performed the right action on 75% of the target trials. Our analysis of looks to the dispreferred item (the animal) found a significant effect of prime only. As predicted, those primed with DO sentences looked more at the animal (Figure 4, right panel). In a secondary analysis we excluded all trials where children committed errors. The effect of prime on looks to the animal persisted. Thus we found that 3-year-olds' interpretation of target dative sentences was influenced by the previous dative sentences that they had heard, even when the prime and target sentences used different verbs and *give* did not appear as a prime or target.

The results reported here demonstrate within- and across-verb priming in both 3- and 4-year-old children. The across-verb priming results can only be explained by representations that are not verb-specific. Therefore, these results suggest that both 3- and 4-year-old children use abstract representations during comprehension. This priming appears across a variety of verbs and in an age group that shows limited productivity in many novel-verb generalization tasks (see Tomasello 2000).

6. Identifying the locus of priming effects in children

What is the source of this priming effect? Our methodology rules out an alternate explanation for previous production priming results found in children. Because the alternate constructions used in the child priming studies are distinguished by the presence or absence of closed-class words or morphemes (*to* for datives or *by* and the participle for passives), production effects could reflect the priming of these words rather than grammatical structures. We avoided this possibility by measuring priming of the role assigned to the direct-object noun, which *precedes* this critical morpheme. This is clearest for the 4-year-olds where our entire time window of analysis preceded the onset of *to* (with a 200 ms offset). However, even for the 3-year-olds, differences between the DO- and PO-prime conditions begin to emerge prior to the onset of this morpheme.

Nevertheless, several alternate explanations of these priming effects remain. The first, and the least interesting, invokes no representation of the phrase structure of the utterance or the argument structure of the verb. Perhaps children in the DO prime conditions simply formed the expectation that the animal would be mentioned first while children in the PO prime conditions learned to expect that the object would be mentioned first. The remaining hypotheses all invoke syntax in one way or another.

The simplest of these hypotheses is that our manipulation directly primed the syntactic structures used in double-object (V NP NP) and prepositional (V NP PP) datives (Pickering & Branigan 1998). These structures would activate the thematic roles associated with them, which in turn would activate animacy features associated with those roles, resulting in the observed eye movements. Alternately, our priming manipulation could have targeted the mapping between thematic roles or animacy features on the one hand and syntactic positions on the other. For example, if the locus of the effect was the mapping of thematic roles, DO primes would potentiate the recipient←→direct object mapping, while PO primes would potentiate a theme←→direct object mapping. Since the recipient and theme roles are in turn correlated with animacy (the recipient is usually animate, theme is usually inanimate), this would give rise to the pattern of eye movements seen in our experiments. Alternatively, direct mappings between animacy features and syntactic positions (e.g., animate←→direct object or inanimate←→direct object) may have been primed. All three of these forms of priming have been found in adults during sentence production (syntactic structures: Bock & Loebell 1990; animacy mappings: Bock et al. 1992; thematic role mappings: Chang, Bock, & Goldberg 2003).

To disentangle these different possibilities we will have to examine a broader range of prime and target types to determine which of these features must overlap for robust structural priming. We have begun exploring this issue with the story priming task. This task has the advantage of allowing us to use prime sentences that are not commands and cannot be acted out. Our current work examines the effects of four kinds of primes (4a-d) on the interpretation of DO and PO datives.

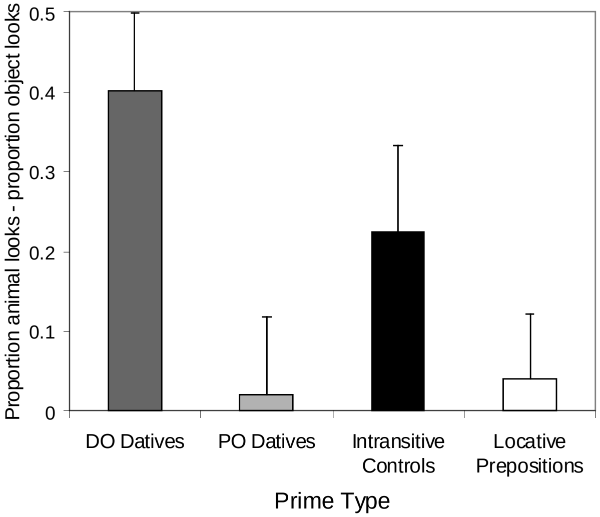


Figure 5. Story priming task with 4-year-olds. Effects of four prime types on the interpretation of dative sentences (PO and DO targets collapsed) as measured by the difference between the proportion of looks to the animate match and the proportion of looks to the inanimate match (higher values are consistent with DO priming).

To date, thirty-two 4-year-olds (eight per prime type) have participated in this study of the nature of between-verb priming.

(4) a. DO prime: She read the girl a story.
b. PO prime: She read a story to the girl.
c. Intransitive prime: She winked.
d. Locative preposition prime: She carried the girl to the bed.

The preliminary results are promising (Figure 5). The observed pattern for the DO and PO primes replicates and extends the results of the act-out priming task. Participants who hear DO-primes show a strong preference for the animal during the region of ambiguity, while those who hear PO-primes have no strong preference for either the animal or the object. Thus dative priming in children persists even when prime sentences are not mapped to an array or an action, and when the prime and target sentences differ along several syntactic and semantic dimensions (e.g., primes in the present study have subjects, include indefinite NPs and in many cases abstract themes).

The locative preposition primes offer a preliminary answer to our questions about the nature of these priming effects. These utterances have the same mapping between animacy and position as the DO datives (animate←→direct object and inanimate←→second NP). However, they have the same syntactic structure as the PO datives (NP PP) and a similar pattern of thematic role assignments (theme, goal/recipient). In this condition the looks during the ambiguous region clearly pattern with the PO primes: the locative prime increases looks to the object relative to the intransitive control resulting in a difference score that is reliably smaller that the DO primes but indistinguishable from the PO primes (see Bock & Loebell 1990 for parallel findings in adult production priming). Thus we tentatively conclude that comprehension priming of datives in preschoolers is not linked to animacy features but may be attributable to the priming of syntactic frames or patterns of thematic role assignment.

7. Conclusions

7.1 Comparison to previous studies

Our results add to the existing literature in several ways. First, they provide converging evidence that young children have abstract structural representations. Our findings complement those of novel-verb comprehension studies. Those studies show that children can generalize attested structures to new verbs in the absence of lexically-specific evidence for these structures. In contrast, our data demonstrate that children use abstract representations in a situation where they know the verbs and could presumably rely solely on lexically-specific representations, were this their dominant form of grammatical representation. Second, these studies demonstrate that these abstractions are active in children during on-line comprehension. Finally, these findings add to the nascent literature on structural priming in preschoolers. Our method rules out some alternate explanations for production priming (e.g., priming of closed-class items), and our results show that abstract structural priming is not restricted to production as suggested by some (Arai et al. 2007).

But how can we reconcile these findings with the results of novel-verb production studies? One possibility would be to extend recent proposals by usage-based theorists that different tasks tap representations of different strength. For example, Tomasello and Abbot-Smith (2002: 212) suggest that "linguistic and other cognitive representations grow in strength during ontogeny, and performance in preferential-looking tasks requires only weak representations whereas performance in tasks requiring more active behavioral decision making requires stronger representations". Perhaps weak abstract representations also suffice to produce the across-verb comprehension priming reported here. We see two reasons to be skeptical of this analysis. First, the pattern of data that we observed provides no evidence for a developmental shift in the relative strength of lexical and abstract representations. As we discuss below, our paradigm allows us to estimate the relative size of abstract priming and lexically-specific priming. In both age groups we find robust evidence for abstract priming in the form of reliable between-verb priming but no reliable evidence of lexically-specific priming (i.e., no interaction between prime type and within-/between-verb priming in either age group). In fact, the evidence for lexically-specific priming is particularly weak in the 3-year-olds, where the difference in effect size between the within-verb and between-verb priming is negligible. Second, the graded strength hypothesis is weakened by a recent study demonstrating that 3-year-olds can generalize the dative alternation in a novel-verb *production* task (Conwell & Demuth 2007). Thus, by 3 years of age, some abstract representations are clearly strong enough to influence both comprehension and production, suggesting the need for an alternate explanation of the co-existence of item-specific use and abstract structural representations. Below, we describe how structural priming can be used to investigate one such alternate explanation which has been widely accepted by those studying adult sentence processing.

7.2 Using structural priming to investigate children's representations

The structural priming technique offers promise for exploring the theoretical and developmental issues raised in the Introduction. Theoretical work on argument structure has consistently acknowledged both broad syntax-semantics correspondences and the role that lexical information plays in the syntactic realization of event structure (Dowty 1991; Levin 1993; Jackendoff 2002). Developmentally, there is a tension between evidence for early abstract representations (e.g., Fisher 2002b) and item-specific use (e.g., Tomasello 2000). Studying the relation between children's lexical representations and their abstract representations may be a fruitful avenue for resolving these questions.

Lexical-specificity and abstract syntax have long been accepted and reconciled in theories of adult sentence comprehension. The data have left us with little choice. For example, Trueswell and Kim (1998) found that reading times for temporarily ambiguous sentence complements like (5) were affected by brief exposures (39 ms) to one-word primes.

(5) The photographer accepted the fire could not be put out.

Exposure to a verb that typically takes a sentence complement (e.g., *realize*) facilitated ambiguity resolution, while exposure to a verb that typically takes a direct object (e.g., *obtain*) hindered it. Note that this effect can only take place in a representational system which is both lexically specific (different verbs had different impacts) and abstract (the structural biases of one verb affected processing of another). Like many in the field, the authors accounted for these findings by positing that individual verbs are associated with abstract structural representations which can be primed (see e.g., MacDonald, Pearlmutter, & Seidenberg 1994). The strength of the link between the verb and a structural node depends on learner's prior exposure to that particular verb in that particular syntactic context.

Models of this kind provide an explanation for the co-existence of generalization and item-specificity in young children (see Fisher 2002b). Perhaps young children have the same linguistic architecture as adults, but simply lack experience. Perhaps like adults, they have abstract syntactic representations, abstract semantic representations, and mappings between the two. However, because their experience with individual verbs is limited the connections between some individual verbs and some structures may be weaker or even absent. Were this true, we would expect children to succeed when knowledge of the construction alone is sufficient to solve the problem. This is generally the case in novel-verb preferential-looking studies. The structure is provided and the use of general linking rules is sufficient to interpret the utterance without integrating verb-specific information. However, when the task requires children to use the connection between the verb and the structure, we would expect that performance would depend upon (1) the child's prior experience with the verb in that structure and (2) their experience of the verb in alternate structures. Novel-verb production studies put the child in precisely the situation where she is least likely to be able to link the verb to the new structure: there is no prior association between the two and there is a strong association between the verb and an alternate structure. Known-verb priming studies allow the child to make use of previously acquired associations between specific verbs and abstract structures.

Similar models have been invoked to explain the differences between within- and across-verb priming in production studies. As we noted above, while structural priming occurs even when utterances share no content words, some researchers have found that priming is greater when the same verb is used in both target and prime (Pickering & Branigan 1998). Pickering and Branigan explain this pattern with a theory in which individual verbs are linked to abstract combinatorial representations such as [NP, NP] and [NP, PP]. These abstract combinatorial nodes are shared between verbs, leading to across-verb priming. In addition, the link between an individual verb and a combinatorial node can be potentiated, leading to an advantage for within-verb priming.

To explore whether our results accord with this pattern, we compared within- and across-verb priming in 3- and 4-year-old children. For 3-year-olds, the effect sizes for within- and across-verb priming were partial $\eta^2 = .21$ and partial $\eta^2 =.17$, respectively. For 4-year-olds, the within-verb and across-verb priming effect sizes were partial η^2 =.45 and partial η^2 =.16 respectively. Thus, for both age groups within-verb priming appears to be stronger than across-verb priming. However, the interaction between the two types of priming was not significant for either group.

7.3 Final words

The studies presented in this chapter demonstrate that the on-line interpretation of dative utterances can be structurally primed by prior comprehension of other dative sentences. This priming effect is robustly present in adults, 4-year-olds, and 3-year-olds and appears regardless of whether the prime sentence is acted out or mapped onto a visual display. These effects are not, or not solely, lexically specific. They persist when different verbs and nouns are used in the prime and target sentences. The between-verb priming effect demonstrates that children as young as 3 years employ abstract representations during the comprehension of sentences with known verbs.

Future studies can shed light on important questions that remain. These include the precise nature of the representations that can be primed (semantic, syntactic, or mappings between syntax and semantics), and the constraints on priming between verbs (is priming restricted to verbs with similar distribution, similar meaning, or both?). Critically, future priming studies can elucidate whether young children like adults, have a language processing system in which lexical and abstract representations interact to produce both item-specific and generalized patterns of use.

Acknowledgments

We thank Cynthia Fisher and Steven Pinker for helpful suggestions and comments. We are grateful to Sneha Rao, Jane Sung, Adrianna Saada, Megan Powell and Alison Young for their assistance in collecting and coding data. This work was funded by a grant from the National Science Foundation (NSF-BCS 0623945) and we thank them for their support. The adult experiments summarized in Section 4 of this paper appear as Experiments 1 and 2 in Thothathiri and Snedeker (in press, a). The experiments with 3- and 4-year-olds from Section 5 are reported in Thothathiri and Snedeker (in press, b). The data reported in Section 6 come from a subset of the participants tested in Thothathiri and Snedeker (2007), Experiment 1.

References

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory & Language, 38*, 419–439.

Arai, M., Van Gompel, R. P. G., & Scheepers, C. (2007). Priming ditransitive structures in comprehension. *Cognitive Psychology, 54*(3), 218–250.

Baker, M. (1988). *Incorporation: A theory of grammatical function changing.* Chicago IL: University of Chicago Press.

Baker, M. (1997). Thematic roles and syntactic structure. In L. Haegeman (Ed.), *Elements of grammar* (pp. 73–137). Dordrecht: Kluwer.

Bock, K. (1986). Syntactic persistence in language production. *Cognitive Psychology, 18*, 355–387.

Bock, K. (1989). Closed class immanence in sentence production. *Cognition, 31*, 163–186.

Bock, K. & Loebell, H. (1990). Framing sentences. *Cognition, 35*, 1–39.

Bock, K., Loebell, H., & Morey, R. (1992). From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological Review, 99*(1), 150–171.

Bowerman, M. (1973). *Early syntactic development: A cross-linguistic study with special reference to Finnish.* Cambridge: CUP.

Braine, M. D. S. (1976). Children's first word combinations. *Monographs of the Society for Research in Child Development, 41*(1, Serial no. 164).

Branigan, H. P., McLean, J. F., & Jones, M. W. (2005). A blue cat or a cat that is blue? Evidence for abstract syntax in young children's noun phrases. In A. Brugos, M. R. Clark-Cotton, & S. Ha (Eds.), *The proceedings of the 29th annual Boston University conference on language development* (pp. 109–121). Somerville MA: Cascadilla.

Branigan, H. P., Pickering, M. J., Stewart, A. J., & McLean, J. F. (2000). Syntactic priming in spoken production: Linguistic and temporal interference. *Memory & Cognition, 28*(8), 1297–1302.

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co-ordination in dialogue. *Cognition, 75*, B13–25.

Branigan, H. P., Pickering, M. J., & McLean, J. F. (2005). Priming prepositional-phrase attachment during comprehension. *Journal of Experimental Psychology: Learning, Memory & Cognition, 31*(3), 468–481.

Campbell, A. L. & Tomasello, M. (2001). The acquisition of English dative constructions. *Applied Psycholinguistics, 22*, 253–267.

Chang, F., Bock, K., & Goldberg, A. E. (2003). Can thematic roles leave traces of their places? *Cognition, 90*, 29–49.

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review, 113*(2), 234–272.

Conwell, E. & Demuth, K. (2007). Early syntactic productivity: Evidence from dative shift. *Cognition, 103*(2),163–179.

Dowty, D. (1991). Thematic proto-roles and argument selection. *Language, 67*(3), 547–619.

Fagan, J. F. (1970). Memory in the infant. *Journal of Experimental Child Psychology, 9*(2), 217–226.

Fantz, R. L. (1961). The origin of form perception. *Scientific American, 204*, 66–72.

Fernald, A., Zangl, R., Thorpe, K., Hurtado, N., & Williams, C. (2006). Learning to listen ahead in English and Spanish: Infants use multiple linguistic and nonlinguistic cues in online sentence interpretation. Paper presented at the Workshop on On-Line Methods in Children's Language Processing, CUNY.

Fillmore, C. J. (1968). The case for case. In E. Bach & R. T. Harms (Eds.), *Universals in linguistic theory* (pp. 1–88). New York NY: Holt, Rinehart and Winston.

Fisher, C. (2002a). Structural limits on verb mapping: The role of abstract structure in 2.5-year-olds' interpretations of novel verbs. *Developmental Science, 5*(1), 55–64.

Fisher, C. (2002b). The role of abstract syntactic knowledge in language acquisition: A reply to Tomasello. *Cognition, 82*, 259–278.

Gamez, P. B., Shimpi, P. M., & Huttenlocher, J. (2005). Emerging syntactic representations as evident through priming. Poster presented at the 2005 SRCD Biennial Meeting, Atlanta, GA.

Gertner, Y., Fisher, C., & Eisengart, J. (2006). Learning words and rules: Abstract knowledge of word order in early sentence comprehension. *Psychological Science, 17*, 684–691.

Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition, 73*, 135–176.

Gleitman, L. R. (1990). The structural sources of word meaning. *Language Acquisition*, *1*, 3–55.

Goldberg, A. (1995). *Constructions: A construction grammar approach to argument structure.* Chicago IL: University of Chicago Press.

Goldberg, A. (2006). *Constructions at work: The nature of generalization in language.* Oxford: OUP.

Gombert, J. E. (1992). *Metalinguistic development.* Chicago IL: University of Chicago Press.

Grimshaw, J. (1990). *Argument structure.* Cambridge MA: The MIT Press.

Grimshaw, J. (1981). Form, function, and the language acquisition device. In C. L. Baker & J. McCarthy (Eds.), *The logical problem of language acquisition* (pp. 165–182). Cambridge MA: The MIT Press.

Gropen, J., Pinker, S., Hollander, M., Goldberg, R., & Wilson, R. (1989). The learnability and acquisition of the dative alternation in English. *Language, 65,* 203–257.

Henderson, J. M. & Ferreira, F. (Eds.). (2004). *The interface of language, vision, and action: Eye movements and the visual world.* New York NY: Psychology Press.

Hirsh-Pasek, K. & Golinkoff, R. M. (1996). *The origins of grammar: Evidence from early language comprehension.* Cambridge MA: The MIT Press.

Huttenlocher, J., Vasilyeva, M., & Shimpi, P. (2004). Syntactic priming in young children. *Journal of Memory and Language, 50*, 182–195.

Jackendoff, R. (2002). *Foundations of language – brain, meaning, grammar, evolution*. Oxford: OUP.

Konopka, A. & Bock, K. (2005). Helping syntax out: How much do words do? Paper presented at the 18th Annual CUNY Sentence Processing Conference, Tucson, AZ.

Levin, B. (1993). *English verb classes and alternations*. Chicago IL: University of Chicago Press.

Luka, B. J. & Barsalou, L. W. (2005). Structural facilitation: Mere exposure effects for grammatical acceptability as evidence for syntactic priming in comprehension. *Journal of Memory and Language, 52*, 436–459.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review, 101*(4), 676–703.

Marslen-Wilson, W. D. & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology, 10*(1), 29–63.

Matin, E., Shao, K. C., & Boff, K. R. (1993). Saccadic overhead: Information-processing time with and without saccades. *Perception & Psychophysics, 53,* 372–380.

Ninio, A. (2005). Testing the role of semantic similarity in syntactic development. *Journal of Child Language, 32,* 35–61.

Ninio, A. (1999). Pathbreaking verbs in syntactic development and the question of prototypical transitivity. *Journal of Child Language, 26,* 619–653.

Noppeney, U. & Price, C. J. (2004). An fMRI study of syntactic adaptation. *Journal of Cognitive Neuroscience, 16*, 702–713.

Pickering, M. J. & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory & Language*, *39*, 633–651.

Pickering, M. J. & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences, 27*(2), 169–226.

Pinker, S. (1984). *Language learnability and language development*. Cambridge MA: Harvard University Press.

Savage, C., Lieven, E., Theakston, A., & Tomasello, M. (2003). Testing the abstractness of children's linguistic representations: Lexical and structural priming of syntactic constructions in young children. *Developmental Science, 6*, 557–567.

Savage, C., Lieven, E., Theakston, A., & Tomasello, M. (2006). Structural priming as implicit learning in language acquisition. *Language Learning & Development, 2*(1), 27–49.

Scheepers, C. & Crocker, M. W. (2004). Constituent order priming from reading to listening: A visual-world study. In M. Carreiras & C. J. Clifton (Eds.), *The on-line study of sentence comprehension: Eyetracking, ERP and beyond* (pp. 167–185). Hove: Psychology Press.

Snedeker, J. & Gleitman, L. R. (2004). Why it is hard to label our concepts. In D. G. Hall & S. R. Waxman (Eds.), *Weaving a lexicon* (pp. 255–293). Cambridge MA: The MIT Press.

Snedeker, J. & Trueswell, J. C. (2004). The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology, 49*, 238–299.

Snedeker, J. & Yuan, S. (in press). Effects of prosodic and lexical constraints on parsing in young children (and adults). *Journal of Memory and Language.*

Song, H. & Fisher, C. (2002). Young children's sensitivity to discourse cues in on-line pronoun interpretation. *Proceedings of the 26th Annual Boston University Conference on Language Development*. Somerville MA: Cascadilla.

Song, H.-J. & Fisher, C. (2004). Syntactic priming in 3-year-old children. Paper presented at the 29th Annual Boston University Conference on Language Development, Boston, MA.

Spelke, E. S. (1979). Perceiving bimodally specified events in infancy. *Developmental Psychology, 15*(6), 626–636.

Swingley, D., Pinto, J. P., & Fernald, A. (1999). Continuous processing in word recognition at 24 months. *Cognition, 71,* 73–108.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*, 1632–1634.

Thothathiri, M. & Snedeker, J. (in press, a). Give and take: Syntactic priming during spoken language comprehension. *Cognition*.

Thothathiri, M. & Snedeker, J. (in press, b). Syntactic priming during language comprehension in three- and four-year-old children. *Journal of Memory and Language*.

Thothathiri, M. & Snedeker, J. (2007). Grammatical priming in preschoolers: A new method for tackling old questions in language acquisition. Manuscript.

Tomasello, M. (1992). *First verbs: A case study in early grammatical development*. Cambridge: CUP.

Tomasello, M. (2000). The item-based nature of children's early syntactic development. *Trends in Cognitive Sciences, 4*, 156–163.

Tomasello, M. & Abbot-Smith, K. (2002). A tale of two theories: Response to Fisher. *Cognition, 83,* 207–214.

Tomasello, M. & Brooks, P. (1998). Young children's earliest transitive and intransitive constructions. *Cognitive Linguistics*, *9*, 379–395.

Townsend, D. J. & Bever, T. G. (2001). *Sentence comprehension*. Cambridge MA: The MIT Press.

Trueswell, J. C. & Kim, A. E. (1998). How to prune a garden-path by nipping it in the bud: Fast-priming of verb argument structures. *Journal of Memory and Language, 39*, 102–123.

Trueswell, J., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition, 73*, 89–134.

Yuan, S., Fisher, C., & Snedeker, J. (2007). Counting the nouns: Simple structural cues to verb meaning. Manuscript in preparation.

CHAPTER 6

Language acquisition research

A peek at the past: A glimpse into the future

Helen Smith Cairns

As the title suggests, this final chapter attempts to place the *Workshop on On-Line Methods in Children's Language Processing* and the papers in this volume in a historical context. First, there is a brief review of the 40-year history of research in language acquisition, including the late arrival of on-line methodology. Then the chapter emphasizes the questions now being addressed in current on-line research. Finally, the conclusion suggests desiderata with respect to future progress in our understanding of the development of both linguistic competence and linguistic performance.

The importance of the *Workshop on On-Line Methods in Children's Language Processing*, and of the present volume, can be best understood by placing the research reported in the historical context of the evolving methodology of research in first language acquisition. The questions addressed by the papers and posters presented at the Workshop and described in the preceding chapters allow us to make some predictions about where acquisition research may take us in the future. The title of this chapter is intended to convey these dual goals.

In 1959 I had my first child. I believed that he talked like I did and that I had taught him. I literally didn't notice the omitted auxiliaries and bound morphemes. When he began to overgeneralize past tenses and plurals, I was horrified and believed that it was because of the bad influence of children at his nursery school whose parents hadn't done as good a job as I had. In 1969 I had my last child. By then we knew the characteristic child-like speech patterns, which we eagerly recorded, and I rejoiced in his overgeneralizations because they showed he was learning rules and producing forms he had never heard.

I tell this story to make two points. First, to illustrate the youth of our field. In 1959 I was not the only one who didn't know anything about child language; no one did. In Brown's *A First Language* (1973) the earliest publications cited were in 1963, by Braine (1963), Menyuk (1963a 1963b), and Brown (Brown & Fraser

1963). Knowledge that was non-existent in 1959 was well-developed in 1969, yet still primitive compared to 2007. We have experienced an explosion of knowledge about child language in four decades. My second point is to illustrate that our beliefs shape our experiences. Lest this sound dangerously post-modern, I assure you that I believe we were much closer to truth in 1969 than in 1959, and in 2007 than 1969. As belief shapes experience, theory drives methodology. In empirical research we decide what we need to experience in order to test our hypotheses, and we develop methodologies in order to do so.

In *Aspects of the theory of syntax*, published in 1965, Chomsky introduced the Competence/Performance distinction. Competence was the speaker-hearer's knowledge of his language; performance, the actual use of language in real time. To quote Chomsky: "Linguistic theory is concerned primarily with an ideal speaker-hearer… who is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and errors in applying his knowledge of the language in actual performance" (p. 3). He goes on to say "To study actual linguistic performance, we must consider the interaction of a variety of factors, of which the underlying competence of the speaker-hearer is only one" (p. 4). Thus, from the beginning there were two conceptions of linguistic performance: idealized performance that would perfectly reflect competence and actual performance taking place in real time.

In the early days our goal was to describe children's grammars (competence), and our primary method was to observe naturalistic speech. Diary studies gave way to more systematic modes of data collection. In his 1970 book *The acquisition of language* David McNeill had a chapter on methodology in which he addressed production, comprehension, and elicited imitation, but production was primary.

The early pioneers, Brown, McNeill, Bloom, Braine, Menyuk, Bowerman and others, learned an enormous amount from analyzing naturalistic speech. They described the amazing phenomenon of Child English. That's the language spoken by everyone who is learning English, when they are between the ages of (roughly) 1 and 3, yet few of those people ever hear it spoken. Child English is a language considerably different from Adult English, yet with highly predictable and well-understood characteristics.

Many valuable analyses of the acquisition of basic sentence structure, negation, movement, function words, later the pro-drop phenomenon, and evidence for parameter setting were based primarily on production data. Field work has contributed valuable information on the acquisition of languages other than English, and CHILDES remains an important resource for acquisitionists. However, in those early days our methodology flowed from our desire to describe the regularities in children's speech, believing (correctly) that what was systematically produced could give us information about the child's underlying grammatical knowledge. Furthermore, we sought common features of speech across children, believing (also correctly) that if many different children produced similar non-adult syntactic and morphological forms, this would point to similar mechanisms of acquisition.

We learned a lot, but we made a lot of mistakes, as well. For instance, we thought that young children do not represent function words in their early lexicons (or functional categories in their syntax) nor bound morphemes in their morphologies. When cross-linguistic work began, we discovered that this absence of bound morphemes is peculiar to lightly inflected languages like English. Children in highly inflected languages, such as Italian, produce bound morphemes in their earliest utterances. Clever behavioral studies, such as those by Shipley, Smith, and Gleitman (1969) and Gerken and McIntosh (1993) have since demonstrated that English-speaking children do, in fact, represent function words and bound morphemes long before they use them in speech. So we learned that children know more than they say, and we realized that we needed to concentrate more on developing viable hypotheses about children's implicit knowledge of language, rather than being satisfied with analyses of speech behavior. This is an example of our beliefs shaping what we choose to experience empirically. We initially believed that all we needed was to describe speech patterns, so we listened carefully. Later we came to believe that underlying linguistic knowledge is imperfectly reflected in speech and we developed increasingly sophisticated (mostly off-line) methods for testing hypotheses about children's underlying grammars.

Simultaneously, linguistic theory became more sophisticated. In the early days of transformational grammar we focused on the application of phrase structure rules and transformational rules. In children's speech we sought to discover their increasingly sophisticated application of the rules of language; a great deal of the early work on child language (speech) attempted to account for basic sentence structure with hypothesized phrase structure and transformational rules (e.g., Brown & Hanlon 1970; Menyuk 1969). As linguistic theory became more technical, it postulated the acquisition of knowledge structures far removed from anything that could be discovered by an analysis of actual speech. Since then, our goal has been to characterize children's grammars and to figure out how they could develop over time based on input consisting of exclusively positive evidence constrained by general principles of acquisition, cognition, and development.

Acquisition research turned from an analysis of individual constructions to acquisition of systems of the grammar. We were able to distinguish between universal principles, which we could assume to be available to the child at birth and language-particular aspects of individual languages, which had to be acquired from linguistic input. Research turned to attempts to demonstrate that from the beginning children obey universal constraints, such as apply to movement operations (Crain 1987) and to acquisition of grammatical modules such as binding (Wexler & Chien 1985; McDaniel, Cairns, & Hsu 1990) and control (McDaniel, Cairns, & Hsu 1990/1991). Acquisitionists' theory construction became much more constrained. If children were found to appear to violate universal principles, then it was essential to develop an account of why this was the case. For instance, pragmatic factors were invoked to account for the apparent violation of Principle B of the binding theory (Montalbetti & Wexler 1985; McDaniel & Maxfield 1992). Limited structure-building abilities were hypothesized to account for apparent violation of control principles (Cairns, McDaniel, Hsu, & Rapp 1994).

The theory that languages vary parametrically provided the basis for an explosion of research investigating grammatical acquisition in a variety of languages. A large literature points to the child's very early (before the age of three) acquisition of a number of language specific properties. In many languages, such as Spanish and Italian, the subject of sentences can be phonetically null; in others, such as English, it must be pronounced. This is known as the pro-drop parameter, and it is apparently set very early (Hyams 1986). There are languages such as German and Dutch in which finite verbs appear in the second position of a sentence, while nonfinite verbs appear at the end. Early learning (Poeppel & Wexler 1993) of this parameter demonstrates not only that children have learned a syntactic requirement of their language, but also that they can discriminate finite from non-finite verbs. Children acquiring all languages with a standard ordering of Subject, Verb, and Object (e.g., SVO in languages such as English and Swedish, SOV in German) learn language-particular word order possibly before they are combining words (Wexler 1999). Since word order follows from a parameter regulating phrasal structure, this demonstrates the acquisition, not of a linear constraint, but of a structural one. In French the placement of the morpheme *pas,* indicating negation, varies according to the finiteness of the verb, appearing before a finite verb and after a non-finite one. The fact that children acquire this distinction around the age of two demonstrates the ability to discriminate between the two types of verb and also knowledge of a movement operation that raises finite, but not non-finite verbs (Pierce 1992). Thus, cross-linguistic investigations of language acquisition demonstrate response at very early ages to the information required by children to set parameters and acquire non-universal, language-specific aspects of their languages.

A major discovery of the cross-linguistic study of acquisition has been that languages spoken by members of non-industrial cultures are not "primitive", as was once believed, but to have all the complexity of the languages of advanced industrial cultures. Arguably the most important contribution in cross-linguistic work has been the demonstration that signed languages are fully human languages, with all the universal and language-specific properties of spoken languages

(Klima & Bellugi 1979). In the not-so-distant past signed languages were thought of as simple gestural systems, not fully formed languages. Now we know that children acquiring signed languages go through the same stages as do speaking children: from babbling to "telegraphic" utterances, to fully formed sentences (Petitto 1994).

All of these investigations of grammatical development focused on the acquisition of an underlying system of grammar, or linguistic competence. Other lines of research investigated linguistic performance.

In the early days we thought of "performance factors", memory limitations, speech errors, false starts, and disfluencies, as masking ideal speaker-hearer speech. However, very early in adult psycholinguistics, under the influence of people like Jerry Fodor, Tom Bever, Merrill Garrett, and George Miller, the other meaning of performance, as the production and comprehension of speech based on the underlying grammar, was taken seriously as a focus of experimental investigation. Early work took too literally elements of linguistic theory as elements of psycholinguistic theory, and enterprises such as the Derivational Theory of Complexity got things off to a rocky start. But we soon learned that theories of performance needed to be independent of linguistic theory, although in some important sense compatible with it. Systematic behavioral responses to newly discovered phenomena like garden-path sentences called for explanations that looked beyond a theory of competence. The grammar is agnostic about whether a temporarily ambiguous structure, such as *Everyone believed John*… should contain a clause boundary between *believed* and *John*, making *John* the subject of an embedded clause, or not, making *John* the object of the verb *believed*. The comprehension system, however, has an overwhelming preference for the latter analysis. Thus, there are language processing mechanisms that, while dependent upon the grammar for a range of possible structures, clearly have parsing preferences that are independent of the grammar. Experimental psycholinguists were interested from the beginning in developing theories to account for adults' processing of sentences in real time because many of the preferences of the comprehension system are theoretically temporary and therefore not measurable with off-line techniques that predominantly measure the product of comprehension. So they developed on-line tasks such as phoneme and word monitoring, probe tasks, and other reaction time measures, which reflected the properties of the mechanisms engaged in language processing as it unfolds over time.

Interest in children's linguistic performance as an object of study (rather than as a vehicle for testing hypotheses about their competence) lagged far behind adult psycholinguistics. I think this was primarily because acquisitionists were focused on grammatical development in children and wanted to figure out how children's grammars progressed from the initial to the adult state. (And we wanted our experiments to be as free as possible from "performance factors", such as effects of memory and task demands.) Adult grammars were the province of linguists, while adult processing was the responsibility of experimental psycholinguists. Thus, attention to children's grammatical development eclipsed interest in their processing of language.

In 1996 MIT Press published *Methods of assessing children's syntax* (McDaniel, McKee, & Cairns 1996), which reviewed a large variety of (mostly) off-line methods used to investigate grammatical development. There were only two chapters dealing with on-line methods, one by Kathy Hirsh-Pasek and Roberta Golinkoff on intermodal preferential looking and one that was a general review of on-line methods by Cecile McKee (McKee 1996). She pointed out the rarity of on-line studies relative to off-line ones, citing a study by Tyler and Marslen-Wilson (1981) as one of the earliest attempts to identify operations in the real-time processing of sentences by children. It was a word monitoring experiment investigating various contextual effects on language processing in 5, 7, and 10-year-old children. (A historical note: This was when the *Journal of Memory and Language* was the *Journal of Verbal Learning and Verbal Behavior.*) In their introduction Tyler and Marslen-Wilson write the following: "In most of the research on the development of language… comprehension is measured at the end of an utterance, rather than as it is being heard…. Such studies… have little to say about the internal structure of processing events…. To address these kinds of issues we have to use tasks which tap the operations involved in real-time or "on-line" sentence processing" (p. 400). Such methods have been only gradually developed over the eleven years after McKee wrote her chapter.

This volume and the Workshop that inspired it is extremely important for our field because it focuses on the "real-time" operations alluded to by Tyler and Marslen-Wilson in 1981, a domain of inquiry that is destined to become an increasingly prominent aspect of child language research. It is particularly timely because on-line methodology has matured sufficiently in the last decade to constitute a robust field in its own right.

Not surprisingly, as research in children's language processing has come into its own, the theoretical issues and empirical questions that it addresses are primarily interested in grammatical development. An interesting aspect of this Workshop and of the work discussed in this volume is that the new questions are directly reflected in the new on-line methods that are being developed. Theory identifies the hypotheses and questions that need to be addressed empirically, and clever people figure out the methods to do that. While off-line tasks can test hypotheses about the status of children's grammars, it is crucial to know whether those grammatical principles are applied as sentences are being processed. An early and important study addressing this question was one by McKee, Nicol, and McDaniel (1993). Using a cross modal priming paradigm, they demonstrated that children apply knowledge of Principle A on-line. Further, they showed that the mysterious tendency of some children to (apparently) disobey Principle B manifests itself during sentence processing. Just those children who disobeyed Principle B in an off-line truth-value judgment task showed activation of illegal antecedents in the on-line task, while those who obeyed Principle B in an off-line task did not. The study by Hirsh-Pasek and Golinkoff (1996) using the intermodal preferential-looking paradigm (which they adapted from Spelke, who developed it in 1979 to study intermodal perception) demonstrated that infants have internalized the standard word order of English and can distinguish transitive from intransitive sentence frames. This result was particularly important because the infants they studied were not yet using combinatorial speech.

A closely related question is: What are the representations constructed by children as they process sentences? Obviously, the construction of a representation is based on grammatical knowledge, so any particular representation is *prima facie* evidence of the existence of that knowledge. For example, experiments reported by Clahsen (this volume) demonstrating that children fill gaps on-line demonstrate not only the existence of empty categories in children's representations, but also the existence in their grammars of movement operations. Snedeker and Thothathiri (this volume) argue that effective cross-verbal priming in their structural priming studies demonstrate that children as young as 3 and 4 years old construct abstract representations of verb and sentence structure. The now well-known preference of children for a goal interpretation of the first prepositional phrase in sentences like *Put the frog on the napkin into the box* (Trueswell, this volume) reflect an initial structural analysis in which the first PP is VP rather than NP attached. It also indicates that children represent the argument structure of verbs in their lexicons.

A number of off-line studies (e.g., Gerken & McIntosh 1993) have demonstrated that children who omit function words in speech include them in sentence representations. Work reported at this Workshop by Kedar (2006) reinforces this finding, but shows that even younger children than those identified by off-line studies can be shown to construct internal representations of function words if sufficiently sensitive measures are used (for Kedar looking latencies in the intermodal preferential-looking paradigm). This work is critical to our understanding of when children begin to represent functional categories in their grammars. The late use of function words had led earlier to the belief that early grammars may lack functional categories (Radford 1990).

Just as work in grammatical development asked questions about when children's grammars become adult-like, much processing work is driven by asking how children's processing operations are similar to and different from those of adults. An excellent example of this line of research comes from the on-line processing of lexical ambiguities. David Swinney (1979), using a cross-modal priming task, demonstrated the initially counter-intuitive finding that upon encountering an ambiguous word in a sentence, adults retrieve all possible meanings, independent of context, then select the contextually appropriate meaning before the end of the sentence. This work was of great importance because for years we had known about the role of context in sentence processing in general and in disambiguation in particular. Swinney identified the locus of the context effect, when it does and does not influence the hearer, and revealed the existence of unconscious mental operations that come into play during sentence processing. Ten years later Swinney and Prather (1989), and, later, Love, Swinney, Bagdasaryan, and Prather (1999) showed that this same effect obtains for children. At this Workshop Khanna, Boland and Cortese (2006) reported studies of the resolution of lexical ambiguity in response to biasing context. Using a cross-modal naming paradigm, they showed that second and third grade children (approximately 8 and 9 years old) are not able to use biasing information in the way adults do, but fourth grade children (10 years old) are. Further, successful use of context was correlated with reading ability.

Other on-line work on lexical access and organization demonstrates that children's lexicons are organized similarly to those of adults. In an eye movement study of cohort effects, reported at this Workshop, Sekerina (2006) showed that preschool children demonstrate the cohort effect, but that developmental progression is in the speed and efficiency with which lexical information is accessed and used.

A related issue is whether, during on-line sentence processing, children use the same kinds of information as adults do. Gibson, Breen, Rozen, and Rohde (2006) demonstrate that relative clauses with object extraction are processed more slowly than are those with subject extraction for both adolescents (age 12–15) and adults. The two groups show similar garden-path effects in sentences with reduced subject relative clauses, effects which are attenuated by the presence of verbs with a high frequency of use as a past participle. The tasks were self-paced reading and listening. The striking difference between the two groups was revealed in the reading task, in which the adults, but not the adolescents, were assisted by plausibility information. Note, however, that in the Gibson et al. (2006) study participants were much older than children in most on-line studies. It is unclear whether self-paced reading could be done with much younger children, given the complexity and length of the materials. Still, the lack of adult-like use of plausibility is even more interesting, given the fact that the people studied were adolescents and might have been expected to exhibit fully adult-like processing strategies. Kidd, Stewart, and Serratrice (2006), in an eye movement study, showed a similar tendency of 5-year-olds to rely more on structural information than on either plausibility or information gained from the visual scene. In a sentence like *Chop the tree with the leaves* children ignore the implausibility of *leaves* as an instrument and rely on the verb bias for VP-attachment to produce the implausible situation of using leaves to chop the tree. Note that this also demonstrates the preference of the children, as well as the adults, to construct representations with the prepositional phrase attached to the VP rather than to the NP.

A crucial processing difference between children and adults is children's inability to revise initial parses by shifting from an initial structural representation to a more appropriate one. Fabrizio, Guasti, and Adani (2006) used a self-paced listening experiment with Italian-speaking 9-year-olds to see whether number agreement on the auxiliary verb would cause them to repair an initial subject relative analysis. Many of the children were unable to do this and, as a consequence, misunderstood the sentences. Fabrizio et al. (2006) suggest that for children structural information is more salient than agreement information, similar to Gibson et al.'s (2006) and Kidd et al.'s (2006) findings about structural information relative to plausibility. In his eye movement experiments, Trueswell (this volume) shows that children above the age of 8 can, as do adults, revise their initial VP attached representation of the first PP in *Put the frog on the napkin into the box* when they hear the second PP, restructuring the parse so that the first PP is NP-attached and the second fulfills the sub-categorical requirements of the verb. Children younger than eight cannot perform this reanalysis, even in the presence of a visual context containing two frogs, one of which is on a napkin.

Explanations for the differences between the processing abilities of children and adults lead on-line researchers to examine the development of non-linguistic cognitive capacities. Interestingly, these are often the "performance factors" we have sought to control in off-line experiments. Khanna, Boland, and Cortese (2006) account for the ability of fourth graders, as distinct from that of second and third graders, to use context for lexical disambiguation by postulating the maturation of inhibition and selection capabilities. They suggest that these capacities are underdeveloped before the age of 8, which is also what Trueswell identifies as the age when successful parsing revisions take place. Trueswell argues that increased revision ability is attributable to the development of cognitive control and executive function. He claims that children must overcome cognitive impulsivity in order to revise initial hypotheses about structure and meaning. Gibson et al. (2006) believe that their finding that plausibility affects listening but not reading is a result not of modality *per se*, but of resource allocation. Several studies presented at this Workshop and discussed in this volume identify the effects of memory span. Clahsen (this volume) reports studies showing that children with low memory span do not show the reactivation of antecedents at gaps demonstrated by children with higher memory spans. They also tend to attach structurally ambiguous relative clauses to the second NP, while high span children prefer the first. Fabrizio et al. (2006) found that children with higher memory spans were more likely than children with shorter memory spans to use agreement information to revise initial structural hypotheses.

Questions about the information used in on-line processing have been extended to cross-linguistic research. In sentences such as *Put the frog on the napkin* English-speaking children prefer VP attachment for the PP presumably because the initial verb is sub-categorized for a locative argument. In Korean sentences of this type, verb information is last, while initial information is case marking on *napkin*, which, though ambiguous, is biased toward the locative. Choi and Trueswell (2006) show that both Korean and English speaking children employ a (by hypothesis) universal strategy of using the first reliable information available in their language, verbal for the English-speaking children, case marking for the Korean children. Further demonstration that children exploit the information available in their language is a series of studies by Fernald and colleagues (Fernald, Zangl, Thorpe, Hurtado, & Williams 2006) showing that Spanish-speaking children as young as 3 use gender marking of the adjective to identify the referent of nouns on timed trials. Post-nominal adjectives in Spanish also facilitate the processing of noun phrases, relative to English pre-nominal forms.

A major contribution of on-line research has been an enhanced understanding of the nature of language disorders, and the promise of early detection of children who are at risk. In a series of ERP studies Friederici (Männel & Friederici this volume) identified a number of measures on which infants at risk for specific language impairment differed from those who were not. At-risk 2-month-old infants failed to discriminate long from short syllables as rapidly as did infants who had no family history of SLI. They also report studies showing that infants with a family history of dyslexia respond to duration changes and consonant changes differently than do infants without such a family history. Retrospective studies demonstrate impaired stress perception and responses to incongruous words in the ERP responses of infants who later have language production deficits. Taken together, these studies suggest that children who are at risk for language disorders process speech input differently than do children who do not develop disorders. On-line methods also promise early identification of children at risk. Both Marchman and Fernald (2006) and Pakulak and Neville (2006) identify low socio-economic status as a major predictor of cognitive and linguistic deficits. The latter even identify differences in brain structure associated with SES differences. Pakulak and Neville report a large study of a variety of intervention techniques with low performing SES children, with promising results. These findings have major public policy implications for early childhood education and intervention.

Van der Lely and Fonteneau's (2006) neurolinguistic work on children with specific language impairment (SLI) who have a specific grammatical impairment has demonstrated a neural substrate specialized for syntactic processing. Her work brings us full circle, from an understanding of language impairment to insight about linguistic functioning in the unimpaired brain.

Several studies presented at this Workshop identified predictors in infancy for language disorders later in life. One, however, demonstrates continuity between language skills in typically developing children from the age of 25-months to 8 years. Marchman and Fernald (2006) conducted a longitudinal study of eye movements in response to picture naming. Those children who responded more quickly and accurately at 25 months demonstrated faster vocabulary growth in subsequent years. Furthermore, in a follow-up study when these children were 8 years old, they were tested on standardized language measures. The response times of the children when they were infants correlated significantly with their performance on language and cognitive tests six years later. This remarkable study demonstrates the continuity of very different language skills over developmental time. The importance of such a finding for our understanding of typical language acquisition cannot be over-estimated.

My own work (Cairns, Waltzman, & Schlisselberg 2004) investigates children's metalinguistic ability to detect the ambiguity of lexically and structurally ambiguous sentences. We are interested not in ambiguity resolution, but in the ability to report that a sentence has two possible meanings. We argue that this ability rests on the lexical and structural processing operations studied on-line: access of multiple meanings of ambiguous words in sentences and the ability to construct structural representations of sentences. In particular, we think that ambiguity detection relies on the ability to reprocess lexical representations and revise structural representations. It is no coincidence that the age of 8, which is crucial in the ability to revise structural analyses on-line, is the age at which children begin to be able to detect structural ambiguity. In order to construct two structural representations (necessary for the perception of the ambiguity of structurally ambiguous sentences) children must escape from what Trueswell calls cognitive impulsivity. We find that ambiguity detection is a massive predictor of reading ability in pre-readers through third graders, and we argue that is because it is just those psycholinguistic processes that are recruited in skilled reading. It would be interesting to investigate whether good early readers have less cognitive impulsivity than do poor readers. It is an ideal outcome for on-line investigations to produce results that elucidate not only on-line processing but also the acquisition of higher level operations, such as metalinguistic skill and reading.

The importance of a variety of methodologies to our understanding of child language cannot be over-estimated. This is because we must have theories of at least three linguistic levels and their acquisition to account for language use. First, we must have a theory of linguistic form and organization (the grammar and lexicon) to define the nature of lexical information and the structures that can be computed during sentence production and comprehension. Such theories are typically tested by a variety of off-line methods. Second, we must have a theory of the processing operations involved in accessing the grammar and lexicon in production and comprehension. Such a theory will include parsing preferences and principles of lexical organization (e.g., frequency) about which the grammar is agnostic. On-line methods, such as those addressed at this Workshop and discussed in this volume, are crucial for testing these theories. Finally, we must have a theory of neural organization and operation to account for how the processing operations are implemented and how knowledge of language is developed and represented. Methods probing brain function and organization, some of which were presented at this Workshop and discussed in this volume, are critical to test these theories. In the spirit of predicting the possible direction of future research, I will suggest where progress needs to be made in each of these areas.

Current theories of the development of linguistic competence assume that the infant begins with innate access to the principles and operations made available by Universal Grammar and, thus, does not need to acquire them through interaction with the environment. Language-particular aspects of morphology and syntax, as well as lexical representations, must be acquired through information available in the speech of the child's community. Advances in learnability theory demonstrate that the information the child receives must be exclusively positive, as opposed to negative, information. That is, the child does not have access to information about which linguistic forms are not available in his language. Note that most of this conception of acquisition theory derives directly from linguistic theory itself. Universal Grammar specifies the innate aspects of language that children do not have to learn, as well as the parameters that must be set through experience. I would like to suggest, however, that we do not really have a theory of language acquisition. As Männel and Friederici state in the first sentence of their paper in this volume: "The wonder of language acquisition with its remarkable speed and its high success remains a mystery." Decades of research have revealed an enormous amount of valuable descriptive information about what children know and when they know it, but we do not yet have a truly explanatory theory. Just as adult psycholinguistics has succeeded in developing processing theories consonant with but independent of the grammar, we need a theory of language acquisition that accounts for how children operate on their linguistic input to create grammars. Similar to adults in language processing, children must engage cognitive processes that are non-linguistic in order to operate on the speech input available in their environment. Slobin (1985) in a series of papers, books, and chapters, described the set of "operating principles" that the child brings to bear on the language learning process. Indeed, we need a theory of language learning universals that all human children employ in order to create an internalized grammar, shaped by biologically-based access to Universal Grammar interacting with the speech of the child's community. Valian (1990) has addressed the question of whether parameters can be set by precise "triggers" in the input language, or whether the child uses parametric variation to construct hypotheses about the language-particular aspects of his language. A discussion of the advantages and difficulties associated with both conceptions would take us too far afield; however, I believe that this is the kind of question we should be addressing as we move toward the goal of a truly explanatory theory of the acquisition of competence.

Similarly, we need a theory of how processing skills develop. Having established that adults use non-linguistic processing operations, we must account for how infants grow up to be adults in this realm, as well as in the realm of linguistic competence. Studies presented at this Workshop suggest that properties of the language, such as word order constraints, movement operations, and lexical information drive early parsing preferences. But how do children acquire processing operations and preferences, such as subject-object asymmetries and the expectation of subject gaps, which are not driven by the language and which the child cannot, in principle, observe? Lexical organization and receptive access appears to be similar for adults and children, but what principles of lexical development can account for this? Can we identify universal principles of structural and lexical processing? How can we account for the fact that some kinds of information seem to be less salient for children than for adults, e.g., information about morphology and plausibility? Non-linguistic cognitive characteristics seem to drive the acquisition of processing to a much greater degree than they do the acquisition of grammar. Selection, inhibition, efficiency of lexical access, cognitive control, executive function, resource allocation, and memory span all seem to play roles in the movement of children to adult-like processing capabilities. We need theories of how these cognitive abilities develop and how they interact with the application of processing operations derived from grammatical and lexical knowledge.

A theory of the development of linguistic performance must address language production as well as reception. We began by acknowledging that children know more than they say: the question is why should that be so. There has been surprisingly little research into the language production of young children, yet there are important cross-linguistic similarities pointing to the role of general linguistic, cognitive, or motoric principles in early speech production. It has been suggested that the well-known "vocabulary spurt" that occurs in the second year of life may be more a result of enhanced lexical access than of word learning (Dapretto & Bjork 2000). A theory of early production will, like theories of adult production, address the mechanisms by which morphological and syntactic forms are realized by the production system, as well as the processes of lexical representation and retrieval (Garrett 1988; Bock & Levelt 1994).

A theory of the development of the neurological representation of language and the brain mechanisms that underlie language learning and processing will undoubtedly be eventually subsumed as a component of developmental neurology. It will be critical to identify those areas of the brain that develop as language learning takes place, just as we now know quite a bit about the time-course of linguistic operations in the developing brain.

As we speculate about the future applications of research in all three areas of language – competence, processing, and the neurological substrate – we are struck by the importance of work in all three to identify children at risk for various types of language disorders. Several presentations at the Workshop and in the wider literature suggest subtle precursors to specific language impairment and other forms of language disorders. Early detection is extremely important, but of equal value is the application of psycholinguistic principles to intervention in disorders of speech, language, and reading. The conversation between psycholinguistic researchers and people on the front lines of helping at-risk and language disordered children must be greatly improved. This will require efforts on both sides of that conversation. Psycholinguists and other language researchers need to reach out to make their findings accessible and relevant to speech-language pathologists and educational specialists. By the same token, practitioners must be willing to listen to people investigating basic language processes and be open to the implications of experimental work for clinical intervention. Enhanced communication across the disciplinary divides could result in improvement in the lives of thousands of children.

Whatever the future holds for the field of language acquisition, we can predict that advances in our knowledge will be as great as they have been in the past. Along with the present volume, the *Workshop on On-line Methods in Childrens' Language Processing* served not only to demonstrate sophisticated new methodologies being developed. It showcased the talented and dedicated scholars who will lead the field forward.

Acknowledgments

I am grateful to Eva Fernández and Dana McDaniel for insightful comments on previous drafts of this chapter.

References

Bock, K. & Levelt, W. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945–978). New York NY: Academic Press.

Braine, M. D. S. (1963). The ontogeny of English phrase structure: The first phase. *Language, 39,* 1–14.

Brown, R. (1973). *A first language: The early stages.* Cambridge MA:Harvard University Press.

Brown, R. & Fraser, C. (1963). The acquisition of syntax. In C. N. Cofer & B. Musgrave (Eds.), *Verbal behavior and learning: Problems and processes* (pp. 158–201). New York NY: McGraw-Hill.

Brown, R. & Hanlon, C. (1970). Derivational complexity and order of acquisition in child speech. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 11–54). New York NY: Wiley.

Cairns, H. S., McDaniel, D., Hsu, J. R., & Rapp, M. (1994). A longitudinal study of principles of control and pronominal reference in child English. *Language, 70,* 260–288.

Cairns, H. S., Waltzman, D., & Schlisselberg, G. (2004). Detecting the ambiguity of sentences: Relationship to early reading skill. *Communication Disorders Quarterly, 25,* 68–78.

Choi, Y. & Trueswell, J. C. (2006). Putting first things last: A cross-linguistic investigation of the developing sentence processing mechanism. Paper presented at the Workshop on On-Line Methods in Children's Language Processing, CUNY.

Chomsky, N. (1965). *Aspects of the theory of syntax.* Cambridge MA: The MIT Press.

Crain, S. (1987). Language acquisition in the absence of experience. *Behavioral and Brain Sciences, 14,* 597–650.

Dapretto, M. & Bjork, E. L. (2000). The development of word retrieval abilities in the second year and its relation to early vocabulary growth. *Child Development, 71,* 635–648.

Fabrizio, A., Guasti, M. T., & Adani, F. (2006). Relative clause processing by Italian children: A self-paced listening study. Paper presented at the Workshop on On-Line Methods in Children's Language Processing, CUNY.

Fernald, A., Zangl, R., Thorpe, K., Hurtado, N., & Williams, C. (2006). Learning to listen ahead in English and Spanish: Infants use multiple linguistic and non-linguistic cues in online sentence interpretation. Paper presented at the Workshop on On-Line Methods in Children's Language Processing, CUNY.

Garrett, M. F. (1988). Processes in language production. In F. J. Newmeyer (Ed.), *Linguistics: The Cambridge survey: Vol. III. Language: Psychological and biological aspects* (pp. 9–96). Cambridge: CUP.

Gerken, L. A. & McIntosh, B. J. (1993). The nterplay of function morphemes and prosody in early language. *Developmental Psychology, 29,* 448–457.

Gibson, E., Breen, M., Rozen, S., & Rohde, D. (2006). Language processing in children as measured using self-paced reading and listening. Paper presented at the Workshop on On-Line Methods in Children's Language Processing, CUNY.

Hirsh-Pasek, K., & Golinkoff, R. M. (1996). The intermodal preferential looking paradigm: A window into emerging language comprehension. In D. McDaniel, C. McKee, & H. S. Cairns (Eds.), *Methods of assessing children's syntax.* Cambridge MA: The MIT Press.

Hyams, N. (1986). *Language acquisition of the theory of parameters.* Boston: Reidel.

Kedar, Y. (2006). Uncovering early grammatical knowledge: Different methods and measures capture specific aspects of infants' linguistic processing competence. Paper presented at the Workshop on On-Line Methods in Children's Language Processing, CUNY.

Khanna, M. M., Boland, J. E., & Cortese, M. J. (2006). Development of sentence context use: When and how do children know that *tag* is a label and not a game? Paper presented at the Workshop on On-Line Methods in Children's Language Processing, CUNY.

Kidd, E., Stewart, A., & Serratrice, L. (2006). Can children overcome lexical biases? The role of the referential scene. Paper presented at the Workshop on On-Line Methods in Children's Language Processing, CUNY.

Klima, E. & Bellugi, U. (1979). *The signs of language*. Cambridge MA: Harvard University Press.

Love, T., Swinney, D., Bagdasaryan, S., & Prather, P. (1999). Real-time processing of lexical ambiguities by children. *Proceedings of the 12th Annual CUNY Conference on Human Sentence Processing.* New York: City University of New York.

Marchman, V. A. & Fernald, A. (2006). Online speech processing efficiency in infancy is related both to vocabulary growth and to school-age language accomplishments. Paper presented at the Workshop on On-Line Methods in Children's Language Processing, CUNY.

McDaniel, D. & Maxfield, T. L. (1992). Principle B and contrastive stress. *Language Acquisition: A Journal of Developmental Linguistics, 2,* 337–358.

McDaniel, D., Cairns, H. S., & Hsu, J. R. (1990). Binding principles in the grammars of young children. *Language Acquisition: A Journal of Developmental Linguistics, 1,* 121–138.

McDaniel, D., Cairns, H. S., & Hsu, J. R. (1990/1991). Control principles in the grammars of young children. *Language Acquisition: A Journal of Developmental Linguistics, 1,* 297–336.

McDaniel, D., McKee, C., & Cairns, H. S. (1996). *Methods for assessing children's syntax.* Cambridge MA: The MIT Press

McKee, C. (1996). On-line methods. In D. McDaniel, C. McKee, & H. S. Cairns (Eds.), *Methods of assessing children's syntax.* Cambridge MA: The MIT Press.

McKee, C., Nicol, J., & McDaniel, D. (1993). Children's application of binding during sentence processing. *Language and Cognitive Processes, 8,* 265–290.

McNeil, D. (1970). *The acquisition of language: The study of developmental psycholinguistics.* New York NY: Harper and Row.

Menyuk, P. (1963a). A preliminary evaluation of grammatical capacity in children. *Journal of Verbal Learning and Verbal Behavior, 2,* 429–439.

Menyuk, P. (1963b). Syntactic structures in the language of children. *Child Development, 34,* 407–422.

Menyuk, P. (1969). *Sentences children use.* Cambridge MA: The MIT Press.

Montalbetti, M. M. & Wexler, K. (1985). Binding is linking. *Proceedings of the West Coast Conference on Formal Linguistics, 4,* 228–245.

Pakulak, E. & Neville, H. (2006). Exploring the relationship between environment, proficiency, and brain organization in children from different socioeconomic backgrounds. Paper presented at the Workshop on On-Line Methods in Children's Language Processing, CUNY.

Pierce, A. (1992). *Language acquisition and syntactic theory: A comparative analysis of French and English child grammars.* Dordrecht: Kluwer.

Petitto, L. A. (1994). Modularity and constraints in early lexical acquisition: Evidence from children's early language and gesture. In P. Bloom (Ed.), *Language acquisition: Core readings* (pp. 95–126). Cambridge MA: The MIT Press.

Poeppel, D. & Wexler, K. (1993). The full competence hypothesis. *Language, 69* (1), 1–33.

Radford, A. (1990). *Syntactic theory and the acquisition of English syntax: The nature of early child grammars of English.* Oxford: Blackwell.

Sekerina, I. A. (2006). Spoken-word recognition in Russian preschoolers. Paper presented at the Workshop on On-Line Methods in Children's Language Processing, CUNY.

Shipley, E. F., Smith, C. S., & Gleitman, L. R. (1969). A study in the acquisition of language: Free responses to commands. *Language, 45,* 322–342.

Slobin, D. I. (1985). Cross-linguistic evidence for the language-making capacity. *The crosslinguistic study of language acquisition,* Vol. 2.: *Theoretical issues* (pp. 1157–1256). Hillsdale NJ: Lawrence Erlbaum Associates.

Spelke, E. (1979). Perceiving bimodally specified events in infancy. *Developmental Psychology, 15,* 626–636.

Swinney, D. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior, 18,* 645–659.

Swinney, D. & Prather, P. (1990). On the comprehension of lexical ambiguity by young children: Investigations into the development of mental modularity. In D. S. Gorfein (Ed.), *Ambiguity processing.* New York NY: Springer.

Tyler, L. K. & Marslen-Wilson, W. D. (1981). Children's processing of spoken language. *Journal of Verbal Learning and Verbal Behavior, 20,* 400–416.

Valian, V. (1990). Null subjects: A problem for parameter-setting models of language acquisition. *Cognition, 35,* 105–122.

Van der Lely, H. & Fonteneau, E. (2006). ERP investigations in typically developing and language-impaired children reveal a domain-specific neural correlate for syntactic dependencies. Paper presented at the Workshop on On-Line Methods in Children's Language Processing, CUNY.

Wexler, K. (1999). Maturation and growth of grammar. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of language acquisition.* New York NY: Academic Press.

Wexler, K. & Chien, Y.-C. (1985). The development of lexical anaphors and pronouns. *Papers and Reports on Child Language Development, 24,* 138–149.

Index

In the series *Language Acquisition and Language Disorders* the following titles have been published thus far or are scheduled for publication:

A complete list of titles in this series can be found on the publishers' website, *www.benjamins.com*

45 **GUIJARRO-FUENTES, Pedro, Maria Pilar LARRAÑAGA and John CLIBBENS (eds.):** First Language Acquisition of Morphology and Syntax. Perspectives across languages and learners. *Expected April 2008*

44 **SEKERINA, Irina A., Eva M. FERNÁNDEZ and Harald CLAHSEN (eds.):** Developmental Psycholinguistics. On-line methods in children's language processing. 2008. xviii, 190 pp.

43 **SAVICKIENĖ, Ineta and Wolfgang U. DRESSLER (eds.):** The Acquisition of Diminutives. A cross-linguistic perspective. 2007. vi, 352 pp.

42 **LEFEBVRE, Claire, Lydia WHITE and Christine JOURDAN (eds.):** L2 Acquisition and Creole Genesis. Dialogues. 2006. viii, 433 pp.

41 **TORRENS, Vincent and Linda ESCOBAR (eds.):** The Acquisition of Syntax in Romance Languages. 2006. viii, 422 pp.

40 **DEEN, Kamil Ud:** The Acquisition of Swahili. 2005. xiv, 241 pp.

39 **UNSWORTH, Sharon, Teresa PARODI, Antonella SORACE and Martha YOUNG-SCHOLTEN (eds.):** Paths of Development in L1 and L2 acquisition. In honor of Bonnie D. Schwartz. 2006. viii, 222 pp.

38 **FRANCESCHINA, Florencia:** Fossilized Second Language Grammars. The acquisition of grammatical gender. 2005. xxiv, 288 pp.

37 **MONTRUL, Silvina A.:** The Acquisition of Spanish. Morphosyntactic development in monolingual and bilingual L1 acquisition and adult L2 acquisition. 2004. xvi, 413 pp.

36 **BARTKE, Susanne and Julia SIEGMÜLLER (eds.):** Williams Syndrome across Languages. 2004. xvi, 385 pp.

35 **SÁNCHEZ, Liliana:** Quechua-Spanish Bilingualism. Interference and convergence in functional categories. 2003. x, 189 pp.

34 **OTA, Mitsuhiko:** The Development of Prosodic Structure in Early Words. Continuity, divergence and change. 2003. xii, 224 pp.

33 **JOSEFSSON, Gunlög, Christer PLATZACK and Gisela HÅKANSSON (eds.):** The Acquisition of Swedish Grammar. 2004. vi, 315 pp.

32 **PRÉVOST, Philippe and Johanne PARADIS (eds.):** The Acquisition of French in Different Contexts. Focus on functional categories. 2004. viii, 384 pp.

31 **MARINIS, Theodoros:** The Acquisition of the DP in Modern Greek. 2003. xiv, 261 pp.

30 **HOUT, Roeland van, Aafke HULK, Folkert KUIKEN and Richard J. TOWELL (eds.):** The Lexicon–Syntax Interface in Second Language Acquisition. 2003. viii, 234 pp.

29 **FERNÁNDEZ, Eva M.:** Bilingual Sentence Processing. Relative clause attachment in English and Spanish. 2003. xx, 294 pp.

28 **SHIMRON, Joseph (ed.):** Language Processing and Acquisition in Languages of Semitic, Root-Based, Morphology. 2003. vi, 394 pp.

27 **SALABERRY, M. Rafael and Yasuhiro SHIRAI (eds.):** The L2 Acquisition of Tense–Aspect Morphology. 2002. x, 489 pp.

26 **SLABAKOVA, Roumyana:** Telicity in the Second Language. 2001. xii, 236 pp.

25 **CARROLL, Susanne E.:** Input and Evidence. The raw material of second language acquisition. 2001. xviii, 461 pp.

24 **WEISSENBORN, Jürgen and Barbara HÖHLE (eds.):** Approaches to Bootstrapping. Phonological, lexical, syntactic and neurophysiological aspects of early language acquisition. Volume 2. 2001. viii, 337 pp.

23 **WEISSENBORN, Jürgen and Barbara HÖHLE (eds.):** Approaches to Bootstrapping. Phonological, lexical, syntactic and neurophysiological aspects of early language acquisition. Volume 1. 2001. xviii, 299 pp.

22 **SCHAEFFER, Jeannette C.:** The Acquisition of Direct Object Scrambling and Clitic Placement. Syntax and pragmatics. 2000. xii, 187 pp.

21 **HERSCHENSOHN, Julia:** The Second Time Around – Minimalism and L2 Acquisition. 2000. xiv, 287 pp.

20 **KANNO, Kazue (ed.):** The Acquisition of Japanese as a Second Language. 1999. xii, 180 pp.

19 **BECK, Maria-Luise (ed.):** Morphology and its Interfaces in Second Language Knowledge. 1998. x, 387 pp.

18 **KLEIN, Elaine C. and Gita MARTOHARDJONO (eds.):** The Development of Second Language Grammars. A generative approach. 1999. vi, 412 pp.

17 **ARCHIBALD, John:** Second Language Phonology. 1998. xii, 313 pp.

16 **HANNAHS, S.J. and Martha YOUNG-SCHOLTEN (eds.):** Focus on Phonological Acquisition. 1997. v, 289 pp.

15 **BRINKMANN, Ursula:** The Locative Alternation in German. Its structure and acquisition. 1997. x, 289 pp.

14 **CLAHSEN, Harald (ed.):** Generative Perspectives on Language Acquisition. Empirical findings, theoretical considerations and crosslinguistic comparisons. 1996. xxviii, 499 pp.

13 **ALLEN, Shanley E.M.:** Aspects of Argument Structure Acquisition in Inuktitut. 1996. xvi, 244 pp.

12 **JUFFS, Alan:** Learnability and the Lexicon. Theories and second language acquisition research. 1996. xi, 277 pp.

11 **YIP, Virginia:** Interlanguage and Learnability. From Chinese to English. 1995. xvi, 247 pp.

10 **LAKSHMANAN, Usha:** Universal Grammar in Child Second Language Acquisition. Null subjects and morphological uniformity. 1994. x, 162 pp.

9 **ADONE, Dany:** The Acquisition of Mauritian Creole. 1994. xii, 167 pp.

8 **HOEKSTRA, Teun and Bonnie D. SCHWARTZ (eds.):** Language Acquisition Studies in Generative Grammar. 1994. xii, 401 pp.

7 **MEISEL, Jürgen M. (ed.):** Bilingual First Language Acquisition. French and German grammatical development. 1994. vi, 282 pp.

6 **THOMAS, Margaret:** Knowledge of Reflexives in a Second Language. 1993. x, 234 pp.

5 **GASS, Susan M. and Larry SELINKER (eds.):** Language Transfer in Language Learning. Revised edition. 1992. x, 236 pp.

4 **ECKMAN, Fred R. (ed.):** Confluence. Linguistics, L2 acquisition and speech pathology. 1993. xvi, 260 pp.

3 **EUBANK, Lynn (ed.):** Point Counterpoint. Universal Grammar in the second language. 1991. x, 439 pp.

2 **HUEBNER, Thom and Charles A. FERGUSON (eds.):** Cross Currents in Second Language Acquisition and Linguistic Theory. 1991. viii, 435 pp.

1 **WHITE, Lydia:** Universal Grammar and Second Language Acquisition. 1989. xii, 198 pp.