Eye-Tracking Technology Applications in Educational Research

Christopher Was
Kent State University, USA

Frank Sansosti
Kent State University, USA

Bradley Morris
Kent State University, USA

A volume in the Advances in Business Information Systems and Analytics (ABISA) Book Series

IGI GLOBAL
DISSEMINATOR OF KNOWLEDGE

www.igi-global.com

Published in the United States of America by
IGI Global
Information Science Reference (an imprint of IGI Global)
701 E. Chocolate Avenue
Hershey PA, USA 17033
Tel: 717-533-8845
Fax: 717-533-8661
E-mail: cust@igi-global.com
Web site: http://www.igi-global.com

Library of Congress Cataloging-in-Publication Data

Names: Was, Christopher, 1967- editor. | Sansosti, Frank J., editor. | Morris, Bradley, 1969- editor.
Title: Eye-tracking technology applications in educational research / Christopher Was, Frank Sansosti, and Bradley Morris, editors.
Description: Hershey PA : Information Science Reference, 2016. | Series: Advances in business information systems and analytics | Includes bibliographical references and index.
Identifiers: LCCN 2016028387| ISBN 9781522510055 (hardcover) | ISBN 9781522510062 (ebook)
Subjects: LCSH: Education--Research--Methodology. | Eye tracking.
Classification: LCC LB1028 .E94 2016 | DDC 370.72--dc23 LC record available at https://lccn.loc.gov/2016028387

This book is published in the IGI Global book series Advances in Business Information Systems and Analytics (ABISA) (ISSN: 2327-3275; eISSN: 2327-3283)

British Cataloguing in Publication Data
A Cataloguing in Publication record for this book is available from the British Library.

Chapter 5
Eye Tracking and Spoken Language Comprehension

Elizabeth Kaplan
Harvard University, USA

Tatyana Levari
Harvard University, USA

Jesse Snedeker
Harvard University, USA

ABSTRACT

Constructing a more precise and deeper understanding of how listeners, and particularly young children, comprehend spoken language is a primary focus for both psycholinguists and educators alike. This chapter highlights how, over the course of the past 20 years, eye tracking has become a crucial and widely used methodology to gain insight into online spoken language comprehension. We address how various eye-tracking paradigms have informed current theories of language comprehension across the processing stream, focusing on lexical discrimination, syntactic analysis, and pragmatic inferences. Additionally, this chapter aims to bridge the gap between psycholinguistic research and educational topics, such as how early linguistic experiences influence later educational outcomes and ways in which eye-tracking methods can provide additional insight into the language processing of children with developmental disorders.

INTRODUCTION

Early developmental research in spoken language comprehension has largely focused on establishing differences between child and adult language abilities; what can adults understand that children cannot? However, for both researchers and educators alike, the central questions lie not just in identifying such differences, but in explaining them. What are the mechanisms that allow adults to succeed where children fail, and how do these mechanisms develop? The difficulty of these questions lies in the nature of language comprehension. Understanding a spoken sentence requires integration of multiple levels of linguistic

DOI: 10.4018/978-1-5225-1005-5.ch005

representation: from phonetic parsing, to identifying lexical entries, determining syntactic structure and semantic relationships, and incorporating discourse context. What's more, adults are able to accomplish this feat seemingly instantaneously, as a sentence unfolds. The advent of eye-tracking methodologies, which use gaze fixations within a visual display as a signal of lexical activation (Allopenna, Magnuson, & Tanenhaus, 1998; Tanenhaus, Spivey-Knowltown, Eberhard, & Sedivy, 1995; Trueswell, Sekerina, Hill, & Logrip, 1999), has allowed researchers to break into the moment-by-moment progression of sentence interpretation, providing a window into the nature of the underlying processing mechanisms.

More recently, these methodologies have been adapted for use with children. Studying children's sentence processing not only provides insight into language acquisition, but also constrains theories of adult language competence. By uncovering how children interpret sentences and the types of errors they make, we can determine the types of information children can and cannot use to understand what they are hearing. Eye-tracking studies have helped to identify the properties of the language system that emerge early in development and track later developing skills as they come online. In the current chapter, we will outline the progress made in the field of psycholinguistics towards specifying the building blocks of the linguistic system, and how it may constrain theories of its underlying structure.

The study of language development also informs educational practices and initiatives for both typically and atypically developing populations. Understanding which sources of information the child language system can use at a particular age can help tailor interventions to be more targeted and effective. Furthermore, a better of understanding of typical language acquisition is necessary to identify where and how breakdowns occur in children with various language deficits. This is imperative in order to create appropriate tools to address the difficulties that such children face. The present chapter aims to outline how progress might be made towards this goal.

BACKGROUND

Cooper (1974) was the first researcher to use an eye-tracking paradigm to explore the relationship between spoken language and eye movements within a constructed visual scene. In his seminal paper, Cooper recorded participants' eye movements to a 3x3 grid containing pictures of different objects as they listened to pre-recorded stories about the displayed items. He found that upon hearing the word for a displayed object, participants rapidly moved their eyes to the corresponding picture. In fact, participants directed their gaze to the correct referent even before the whole word was spoken (Cooper, 1974). The pattern of data suggested that direction of gaze was closely tied to the unfolding linguistic information.

There are two main advantages to using eye-tracking paradigms in psycholinguistic research. First, eye-tracking paradigms allow for a continuous, moment-to-moment measure of spoken language comprehension, whereas offline judgment tasks show only the final interpretation. The fine-grained temporal resolution in these tasks allows for researches to analyze language comprehension on the scale of milliseconds (Allopenna, Magnuson, & Tanenhaus, 1998; Dahan, Magnuson, & Tanenhaus, 2001; Dahan, Magnuson, Tanenhaus, & Hogan, 2001). This is a critical feature because language processing, in both adults and children, is incredibly rapid and incremental; a sentence, and each individual word within it, is processed as it unfolds and before all of the information is available (Snedeker, 2009). Second, eye-tracking paradigms are ideal for use with young children, allowing for researchers to explore previously untestable hypotheses about language acquisition and development. Prior to eye-tracking paradigms, most studies of language focused on written language comprehension because text was simpler to present and could be easily linked to processing speed. Reading tasks, however, pose obvious limitations for studying comprehension in children. In addition, eye-tracking tasks are naturalistic and do not place additional demands on the participant, making data interpretation more externally valid.

Before delving into more detail about how eye-tracking paradigms have been used inform theories of spoken language comprehension, it is important to first establish the connection between the behavioral measure, eye movements, and the underlying cognitive process, language comprehension. Listeners can understand a sentence without any visual context, so why would eye gaze be indicative of linguistic processing? First, eye gaze is a reliable indicator of attention within a visual field (Liversedge & Findlay, 2001). Importantly, when listening to instructions, attention is often directed according to internals goals and thus towards objects that are task-relevant and not just physically salient (Trueswell, 2008). Eye-tracking paradigms use this feature of attention to infer linguistic decisions; a listener's attention is guided by the unfolding sentence and their gaze is automatically directed to objects referenced in the speech stream. The listener's online referential decisions indicate their understanding of the sentence at each moment in time (Trueswell, 2008). (For a more in-depth discussion of the linking hypothesis see Tanenhaus, Magnuson, Dahan, & Chambers, 2000.)

A linking hypothesis between eye-movements and spoken language comprehension is perhaps most explicitly articulated in studies of lexical activation such as Allopenna et al. (1998), which examine how the presence of cohort referents (i.e., words that begin with the same initial consonant and vowel phonemes such as beaker and beetle) and rhyme referents (i.e., words that end with similar phonemes such as beaker and speaker) affect the time course of lexical activation. For this task, participants' eye movements were recorded as they were looking at a computer screen and were instructed to move one of four images: a target image (e.g., beaker), a cohort image (e.g., beetle), a rhyme image (e.g., speaker), and a distractor image (e.g., carriage). Confirming the linking hypothesis, the participants' fixations to the images correlated with the time course of the unfolding acoustic information. Initial fixations to the cohort images early in the target word (at approximately 200ms, when the phonetic input is ambiguous) indicate that the cohort words compete for activation. In other words, during the initial phonemes of the auditory stimulus, participants ruled out distractor images (e.g., carriage) but were just as likely to look at the cohort competitor (e.g., beetle) as they were to the target referent (e.g., beaker). In addition, the rhyme competitor (e.g., speaker) also competed for lexical activation, though not to the same extent as the cohort. Competition from the rhyme indicates activation of not just the incoming phonological properties, but features of whole lexical entries.

The close temporal link demonstrated between lexical activation and corresponding gaze fixation allows researchers to track the progression of linguistic computations as they occur. Though the specific details of the computational levels and how they interact are still widely debated, there is a general consensus about the main processing steps. In the broadest scope, the conceptualization of the language system that we will be working with in this chapter assumes that information is passed from phonological parsing, up to lexical computation and then to syntactic and semantic analysis. Various higher-order cues, like pragmatics, are then integrated into the sentence interpretation. As a sentence unfolds, the incoming information is propagated across these levels incrementally; information is passed to higher levels of processing as it becomes available and often before processing is complete at lower levels. In addition, these levels of analysis mutually constrain each other. Not only is information processed bottom-up, but information from higher levels also influences processing at lower levels.

In the following sections, we will discuss how eye-tracking technology has been used to characterize the mechanisms involved in language comprehension across the processing stream, from lexical discrimination, to syntactic parsing, to higher-order information, specifically pragmatics.

SPOKEN LANGUAGE COMPREHENSION

Lexical Discrimination

Understanding a sentence requires a listener to parse and identify specific words from the acoustic information in the speech stream. Eye-tracking studies with adults have shown that listeners solve this task with remarkable efficiency. Lexical discrimination is incremental; as soon as phonological information becomes available, it is passed up the processing stream and continuously updates activation at the lexical level until a word is settled upon (Marslen-Wilson, & Zwitserlood, 1989; Fernald, Perfors, & Marchman, 2006). Although adults' ability to efficiently discriminate words has been well established, over the past decade and a half eye-tracking research with children has given insight into how this ability develops. The growing consensus is that children's ability to identify words is quantitatively, but not qualitatively, different from that of adults.

Eye-tracking studies, using the 'looking while listening' paradigm have shown that, like adults, children process information as it unfolds. In this paradigm, children are shown a display containing several images and their gaze is monitored while one of the images is named (e.g., Where's the [Target]? Can you see the [Target]?) (Fernald, Swingley & Pinto, 2001). Toddlers as young as 18 months are able to correctly direct their gaze to a target image in a visual display, even when only part of the word is uttered (Fernald, Swingley & Pinto, 2001). The ability to identify a target referent following only the first 300ms of acoustic information signifies that children are not waiting until all the necessary information is revealed before sending the input to higher-order processing at the lexical level. Rather, as information becomes available it continuously constrains possible alternatives (Fernald, Swingley & Pinto, 2001). Although children's lexical processing is qualitatively similar to that of adults, the efficiency with which children can process acoustic information changes across the lifespan. Children across the 2nd year of life show a marked increase in the speed at which they can correctly identify the referent for the spoken target word (Fernald, Pinto, Swingley, Weinberg, & McRoberts, 1998; Fernald, Perfors, & Marchman, 2006).

Studies have shown that the efficiency with which children can direct their gaze to a proper referent in the visual display is related to their rate of vocabulary growth and is predictive of both language and cognitive development during the school-age years. In the Fernald, Swingley, and Pinto (2001) article, described above, children with higher vocabularies were also faster at properly directing their gaze to the target referents, showing a relationship between language production, as indexed by productive vocabulary, and comprehension, as indexed by efficiency of on-line language comprehension. Further, longitudinal studies have shown that the speed at which 25-month olds can identify the target image is related to their concurrent vocabulary, and is also indicative of their vocabulary growth rates across the 2nd year (Fernald, Perfors, & Marchman, 2006). Even more strikingly, in a follow-up study, the same children who were tested at 25 months were brought back into the lab at age 8 and were given standardized measures of language and cognitive development (Marchman & Fernald, 2008). Children who at age 2 had larger vocabularies and faster reaction times in the 'looking while listening' paradigm showed better expressive language scores, while children with low vocabularies and slow reaction times by school-age showed both lower language and IQ scores. Importantly, vocabulary and speed of processing showed independent contributions to performance later in life (Marchman & Fernald, 2008).

The remarkable continuity in language development across childhood needs to be taken into consideration for educational programs and intervention strategies. Firstly, the ability to use eye-tracking technologies to identify a potential delay as early as age two may allow for earlier intervention strategies. If we can identify children who may require assistance earlier, we have a better opportunity to improve their performance prior to school. In addition, the results of the eye-tracking studies reported above suggest that both vocabulary and the efficiency of the language processing system early in childhood play an important role in the construction of the language system and are related to later measures of cognitive performance (Fernald, Perfors, & Marchman, 2006; Marchman, & Fernald, 2008). Thus, both properties of language development should be considered when interventions are designed.

Syntactic Parsing

Following word recognition, sentence comprehension requires parsing of the syntactic structure and identification of the relevant semantic relationships. The adult and child systems for understanding sentence structure are remarkably similar. Both adults and children integrate multiple sources of information in order to predict the most likely syntactic structure based on the current input. However, children through the age of 10 have difficulty using particular cues to inform sentence comprehension (Hurewitz, Brown-Schmidt, Thorpe, Gleitman, & Trueswell, 2000; Kidd & Bavin, 2005; Kidd, Stewart, & Serratrice, 2011; Snedeker & Trueswell, 2004; Trueswell, Sekerina, Hill, & Logrip, 1999; Weighall, 2008). Eye tracking has been invaluable in helping to uncover the types of information that children are and are not able to use when making syntactic decisions, providing insight into the organization of the language processing system and its developmental trajectory.

Differences between child and adult language comprehension have been widely investigated by looking at the processing of ambiguities, such as the prepositional ambiguity in sentence 1.

1. Put the frog on the napkin in the box.

In such sentences, the 1st prepositional phrase, *on the napkin*, is temporarily ambiguous; it can be interpreted as either the goal of the verb, *put*, (VP-attachment; *Put the frog onto the napkin...*) or as a modifier of noun-phrase, *frog,* (NP-attachment; *Put the frog that's on the napkin...*). Using an eye-tracking methodology called a 'visual world paradigm,' Trueswell, Sekerina, Hill, and Logrip (1999) showed participants a set of toys which contained either one or two frogs (1-Referent or 2-Referent conditions) as well as an empty napkin and an empty box.

The study showed that both adults and children are sensitive to the lexical bias of the verb when making syntactic commitments; listeners will originally interpret *on the napkin* as the goal of the verb *put,* since *put* often requires a destination. Critically, adults do not make such a commitment when the visual display contains two frogs, with one of the frogs appearing on a napkin. Instead, they use the referential context to commit to a modifier interpretation of the same prepositional phrase, telling them which frog should be moved. Children, on the other hand, fail to use referential context and commit to a goal interpretation of the prepositional phrase regardless of the number of frogs in the visual scene.

The results of this study, which have been replicated several times (Hurewitz et al., 2000; Weighall, 2008), reveal interesting similarities in child and adult sentence comprehension. Follow-up work using similar paradigms has provided convergent evidence that children are sensitive to various bottom-up sources of information. Like adults, children as young as 4-years old have been shown to use lexical information in order to predicatively commit to syntactic structure (Kidd & Bavin, 2005; Snedeker & Trueswell, 2004; Trueswell et al., 1999). Similarly, children have are able to use prosodic cues to inform syntactic parsing (Snedeker & Yuan, 2008). In a study by Snedeker and Yuan (2008), children were shown a display containing a set of relevant animals and objects as they heard syntactically ambiguous instructions such as sentence 2.

2. You can feel the frog with the feather.

The sentences were presented with prosodic cues that encouraged either a modifier (3) or an instrumental (4) interpretation of the prepositional phrase *with the feather*.

3. You can feel… the frog with the feather.
4. You can feel the frog… with the feather.

Children, like adults, showed significantly different gaze patterns in the two prosodic conditions. When hearing prosodic breaks consistent with a modifier interpretation (3), children directed their gaze to a frog holding a little feather. On the other hand, when hearing prosodic breaks consistent with an instrument interpretation (4), children directed their gaze to a large feather that was more suitable as a tool. These gaze patterns indicate that children are able to use suprasegmental properties of the incoming acoustic stream in order to constrain their understanding of the sentence.

On the other hand, children have difficulty using top-down information to guide parsing decisions (Hurewitz et al., 2000; Kidd & Bavin, 2005; Kidd, Stewart, & Serratrice, 2011; Snedeker & Trueswell, 2004; Trueswell et al., 1999; Weighall, 2008). Not only are children unable to constrain their interpretation based on referential context (Trueswell et al., 1999), but children also struggle to use other, more reliable cues, such as plausibility. Snedeker, Worek, and Shafto (2009) asked children to listen to sentences that had a prepositional phrase that contained either a plausible (5) or an implausible (6) tool for the relevant verb.

5. Tickle the bear with the paintbrush.
6. Tickle the bear with the mirror.

The display shown to the children contained a set of objects that included a bear holding a small mirror, and two possible instruments, a paintbrush and a mirror. Although a mirror is a remarkably unlikely tool for tickling, children seem unable to use this knowledge in order to inform their syntactic commitments. According to lexical bias, they direct their gaze to the instrument mirror, rather than accept a modifier interpretation of the prepositional phrase. In fact, children, more often than adults, carried out actions consistent with such an interpretation and used an implausible tool to carry out the instructions (Snedeker, Worek, & Shafto, 2009). However, although children did make more errors, their offline behavior was more sensitive to global plausibility than indicated in the online measures. The difference between the findings using eye-tracking and the offline actions indicates that children have the ability to use top-down cues, such as plausibility, but that the processing stream is not efficient enough for such higher order information to influence online sentence interpretation (Snedeker, Worek, & Shafto, 2009).

By tracking the ability to resolve linguistic ambiguities using various sources of information, we can begin to more accurately specify the flow of information through the language processing system. However, further research is needed in order to establish the reason for why the use of top-down cues is late developing, and whether this ability reflects language experience or limitations in more general cognitive mechanisms.

Pragmatics

Spoken language is inherently social in nature. Interlocutors exchange information through conversation, yet the information conveyed is often much richer than the syntax and semantics of the utterance alone. How listeners comprehend language is powerfully shaped by inferences about speaker's intentions, the previous discourse, and the manner in which the sentence is spoken. These sorts of inferences, as well as other high-order language abilities, are known as pragmatics. Two lines of research within this field that have particularly benefitted from eye-tracking technology are computation of scalar implicatures and use of prosodic cues.

Scalar Implicatures

Scalar implicatures refer to the pragmatic inferences made when a speaker uses a weaker phrase rather than a stronger alternative expression. The assumption is that speakers in a conversation provide all of the necessary information without being overly informative (Grice, 1975). One of the most common examples of scalar implicatures is the use of the word *some*, as in sentence 7.

7. Billy ate some of the cookies.

The strictly semantic interpretation of sentence 7 does not rule out the possibility that Billy ate all of the cookies, as the logical meaning of *some* has only a lower bound, meaning more than none. However, when most people encounter this sentence, they assume that Billy ate at least one, but not all of the cookies. This inference is based largely on the assumption that if the speaker wanted to communicate that Billy ate all of the cookies, he or she would have used the stronger expression *all* instead of *some*.

Psycholinguistic research has employed eye-tracking methods to explore when and how interlocutors make these sorts of pragmatic inferences. Are listeners able to immediately access the pragmatic interpretation of the word *some*, or is the process incremental, requiring the listener to establish the semantic interpretation prior to formulating the pragmatic inference? Eye-tracking methods have clear advantages for disentangling the relationship between semantics and pragmatics because of their fine-grained temporal measurements.

Recent work using eye tracking to investigate the processing of scalar implicatures has yielded conflicting results. In Huang and Snedeker (2009) participants' eye gaze was measured as they listened to sentences such as sentence 8.

8. *Point to the girl that has* some/all/two/three *of the* socks/soccer *balls.*

Visual displays contained a set of characters (e.g., *the girls*) such that one character had a subset of items (e.g., *socks*) and the other character had a full set of items (e.g., *soccer balls*). In these trials, there was a temporary phonological ambiguity (e.g., *socks* vs. *soccer balls*). Under a strictly semantic interpretation of the critical sentence containing the scalar quantifier *some*, both the girl with *some* and the girl with *all* of the items are equally plausible referents until the disambiguating syllable. However, if the pragmatic inference is available immediately, only the girl who has *some,* but not all, of a set of items is a plausible candidate, allowing subjects to identify the proper referent soon after the onset of the quantifier. Huang and Snedeker also include *two*, and *three* trials to compare to the critical trials.

Participants correctly looked at the target referent during the quantifier phase, before phonological disambiguation, in the *all*, *two*, and *three* trials, but did not do so in the *some* trials. This data pattern suggested that the quantifier *some* was insufficient for disambiguation. There was a delay in reference resolution, suggesting that the pragmatic inference was made after the semantic interpretation was calculated.

However, the relationship between semantics and pragmatics has yet to be resolved. In fact, using a very similar method to Huang and Snedeker's (2009) design, Grodner, Klein, Carbary, and Tanenhaus (2010) came to the opposite conclusion. Their data show that during the *some* trials, subjects looked to the target referent prior to disambiguation, similar to *all* or *none* trials, which do not require a pragmatic inference. Thus, calculating the pragmatic *some* was not delayed relative to the semantic interpretation. Grodner et al. (2010) point out some key differences between their design and that of Huang and Snedeker (2009) that may have caused the disparity in their results. For example, they began each trial with statements about the objects in each set (e.g., There are four balls, four planets and four balloons). In addition to making sure the participant knew the labels for each object, this instruction explicitly drew participants' attention to the full set of each item, potentially making the distinction between full sets and subsets of items more salient. Other differences include shortening the quantifier phase in their instructions and getting rid of exact number trials. Further research is required to fully understand when and under what conditions scalar implicatures are calculated.

Notably, both of these eye-tracking studies are focused on exploring the time course of pragmatic inferences of scalar implicatures in adults. Many questions about the development of the relationship between semantics and pragmatics have yet to be explored. For example, at what age do children make the pragmatic inference that *some* means *some, and not all*? Current research has begun to use eye-tracking methods to investigate the developmental trajectory of scalar implicatures, and how general cognitive mechanisms, such as executive function, play a role.

Prosody

A second line of research within the field of pragmatics that has benefitted greatly from eye-tracking methodologies is the use of prosody as a pragmatic cue. Prosody refers to the patterns of stress and intonation in language, such as pitch variation, volume, and duration. In addition to the semantics of a sentence, the suprasegmental features can often influence a listener's interpretation. Similar to scalar implicatures, questions about prosody often explore the relationship between pragmatic cues and other aspects of language including syntax and semantics. When are listeners able to integrate prosodic cues into their sentence comprehension, and how might these cues affect the ultimate interpretation?

One way in which adults use prosody is to help them make predictions about the unfolding sentence within the context of the previous discourse. For example, it is usually assumed that an accented word refers to something that is new to the discourse, whereas an unaccented word refers to something that is given, or has been previously mentioned (Brown, 1987). Dahan, Tanenhaus, and Chambers (2002) used a visual world paradigm to investigate how quickly adults are able to use this assumption to guide sentence interpretation. In Experiment 1, they monitored participants' eye gaze as they listed to pairs of instructions to move items on a computer screen such as sentence 9.

9. Put the candy below the triangle. Now put the CANDLE/CANDY above the square.

The referent in the critical second instruction would be either previously mentioned (*candy*) or new (*candle*), and would be either accented or unaccented. Similar to the design in Allopenna et al. (1998), this experiment took advantage of the initial phonetic overlap of cohort words (e.g., *candy* and *candle*), which made the referent in the instructions temporarily ambiguous. If prosodic accents had no effect on sentence processing, the proportion of looks between the candy and candle should be approximately equal up until the point of disambiguation; however, this was not the case. Dahan et al. found that participants were significantly more likely to look at the new referent when the target word was accented, and significantly more likely to look at the previously mentioned referent when the target word was unaccented. Moreover, they found that these effects appear right after the onset of the target word, indicating that accent had an immediate effect on reference resolution.

Further support for the idea that prosodic features play an immediate role in adult listeners' online predictions about the upcoming discourse has been illustrated using 'visual search' tasks (Ito, Bibyk, Wagner, & Speer, 2014; Ito, Jincho, Minai, Yamane, & Mazuka, 2012; Ito & Speer, 2008). In this paradigm, participants' eye movements are monitored as they are asked to select objects from an array of items. The items in the array vary in color and are grouped by type (e.g., cats, lions, and monkeys).

Participants are given pairs of instructions to manipulate items in the display, such as sentence 10.

10. Point to the green cat. Now point to the PINK lion/ pink LION.

Repeated results have shown that adults make more and faster anticipatory eye movements to a new target item when there is a felicitous accent on the adjective than when there is an infelicitous accent on the noun of the second instruction (Ito, Bibyk, Wagner, & Speer, 2014; Ito, Jincho, Minai, Yamane, & Mazuka, 2012; Ito & Speer, 2008).

When does the ability to use prosodic accents to make predictions about the upcoming discourse develop? Recent developmental psycholinguistic research has begun to investigate children's use of prosody in reference resolution using similar methods to the adult studies. However, when interpreting eye-gaze data in children, several potential confounds are important to address. These include differences in executive functioning, increased time to prepare a saccade movement compared to adults (Arnold, 2008; Snedeker & Yuan, 2008), and the tendency to perseverate on initial linguistic interpretations (Choi & Trueswell, 2010; Snedeker & Yuan, 2008; Trueswell, Sekerina, Hill, & Logrip, 1999).

Data from both visual world and visual search studies with children have supported the hypothesis that children as young as four- to five-years old are sensitive to the suprasegmental features of language, and are able to use prosody to make pragmatic predictions about the upcoming discourse (Arnold, 2008; Ito, Bibyk, Wagner, & Speer, 2014; Ito, Jincho, Minai, Yamane, & Mazuka, 2012). Compared to adults, however, children are significantly slower in making anticipatory eye movements to the target referents, indicating that the ability to interpret prosody as a pragmatic cue develops over time (Arnold, 2008; Ito, Bibyk, Wagner, & Speer, 2014; Ito, Jincho, Minai, Yamane, & Mazuka, 2012). Ito et al. (2014) explored this developmental trajectory in greater detail using eye-tracking technology during a visual search task in six- to eleven-year old children. They found that with age, there were gradual improvements in how quickly children looked to the target referent after hearing the felicitous accent on the adjective of the second instruction (e.g., Point to the green cat. Now point to the PINK cat).

Suprasegmental cues in the incoming speech stream influence not only the rate at which listeners can identify referents, but also their ultimate interpretation of the sentence, as in Snedeker and Yuan (2008). Though pragmatics and other sorts of higher-order information are typically assumed to be one of the latest stages of sentence processing, they can influence lower levels of processing and play crucial roles in successful language comprehension. Current research focusing on developmentally delayed populations, such as children with Autism Spectrum Disorder, is investigating whether such populations may have difficulty in using these higher-order cues to inform sentence comprehension.

Atypical Language Development

A more complete understanding of the developmental trajectory of spoken language comprehension in typically developing children has also informed our understanding of mechanisms impaired in children with atypical language development. Identifying the stages of language comprehension is essential for pinpointing where along the processing stream abnormal development diverges. Having a more in-depth understanding of these impairments will help researchers, clinicians, and educators structure better treatments and therapies for children struggling with language impairments. Eye-tracking paradigms are particularly useful in studies of developmentally delayed children because the methods remove many of the metalinguistic and executive functioning challenges posed in other methods that could mask or influence the child's linguistic abilities (Nation, 2008). Eye-gaze studies have helped define atypical language processes ranging from low-level linguistic processing speed, as seen in Specific Language Impairment, to higher-order pragmatic deficits, as seen in Autism Spectrum Disorder.

Specific Language Impairment (SLI)

Specific Language Impairment (SLI) is a disorder characterized by language delays, specifically in verbal abilities, in the absence of other nonverbal cognitive deficits (Bishop, 1997) and occurs in approximately 6-7% of kindergarten children in the U.S. (Tomblin, Records, Buckwalter, Zhang, Smith, & O'Brien, 1997). In order to better understand SLI, researchers investigate where along the language comprehension stream these children might begin to have atypical processing. Eye-tracking methodologies, similar to the ones used with adults and typically developing children, have begun to explore this and create a more nuanced understanding of the deficits associated with SLI.

Children with SLI exhibit differences compared to typically developing peers at the lowest level of linguistic processing discussed in this chapter, lexical discrimination. Using a visual world paradigm and stimuli modeled off of Allopenna et al. (1998), McMurray, Samelson, Lee, & Tomblin (2010) investigated differences in online word recognition between typically developing (TD) children, children with SLI, children with non-specific language impairment (NLI), and children with specific cognitive impairment (SCI). Participants listened to spoken instructions to manipulate the visual world, and their eye-movements were recorded as they looked at target items (e.g., *candle*), cohort items (e.g., *candy*), rhyme items (e.g., *handle*), and distractor items (e.g., *button*). McMurray et al. (2010) found that participants with lower language scores, children diagnosed with either SLI or NLI, were significantly slower to show eye movements to the target referent than those without language impairments, TD children or children with SCI. This finding suggests that children with below-average language abilities are significantly slower at lexical discrimination. In fact, children with SLI show eye-gaze patterns similar to those of much younger children who have similar language abilities measured by mean length of utterance (Andreu, Sanz-Torrent, & Guardia-Olmos, 2012). Considering that lexical discrimination is one of the lowest levels of language processing, it is likely that this initial delay becomes exacerbated as information is propagated up through the subsequent levels of the language processing system. Further research in this field is necessary to determine the extent of the delay and its consequences both for sentence comprehension and the development of the language processing system.

Autism Spectrum Disorders (ASD)

Another population that, by definition, has impairments in communication abilities is children diagnosed with Autism Spectrum Disorder (ASD) (APA, 2013). Children across the autism spectrum exhibit wide heterogeneity in their language skills, ranging from completely non-verbal, to having intact articulation, vocabulary and syntax, but all show a marked deficit at social uses of language (Kjelgaard & Tager-Flusberg, 2001; Joseph, Tager-Flusberg, & Lord, 2002).

One of the first studies to explore language-mediated eye movements in children with ASD examined if such children have more difficulty integrating different types of information to inform their language comprehension (Brock, Norbury, Einav, & Nation, 2008). Based on the weak central coherence account of autism, which posits a deficit in integration abilities (Frith, 1989; Minshew, Goldstein, & Siegel, 1997), Brock et al. (2008) predicted that participants on the autism spectrum would have difficulty using cues across levels of linguistic representation. In this study, TD and ASD children were given sentences such as sentence 11 while looking at a display containing both a target noun (e.g., *bucket*) and a phonological competitor (e.g., *buckle*).

11. Joe filled the bucket carefully

They predicted that children with ASD would have difficulty using the lexical bias of the verb in order to constrain possible alternatives and predict the most likely target prior to disambiguation. Their prediction, however, was unsupported by their data and, instead, the researchers found that children's ability to use such information, as demonstrated by their early eye movements to the target referent, was significantly influenced by language scores, not ASD diagnosis.

However, even children with ASD who are highly verbal still exhibit impairments in pragmatics (Tager-Flusberg, Paul, & Lord, 2005; Young, Diehl, Morris, Hyman, & Bennetto, 2005), as evidenced by difficulty using prosody as a pragmatic cue. Diehl, Friedberg, Paul and Snedeker (in press) explored the ability of children with ASD to use prosody to parse syntactically ambiguous sentences such as the ones in Snedeker and Yuan (2008). Using an act out task on a visual world paradigm, Diehl et al. recorded participants' eye-gaze as they listened to syntactically ambiguous instructions to manipulate stimuli in the visual world such as sentence 12.

12. You can pinch the bear with the barrette.

As in Snedeker and Yuan (2008), the prosodic breaks in the instructions were intended to clarify the syntactic ambiguity. The stimuli were blocked so that in the first part of the study, half of the participants received the instrument prosody (13) and the other half of participants received the modifier prosody (14).

13. You can pinch the bear…with the barrette.
14. You can pinch…the bear with the barrette.

In the second part of the study, the conditions were flipped. This block design allowed for researchers to compare across diagnostic groups for the first part of the study, and within diagnostic groups for the second. Both the offline results from the act out task and eye-gaze data from the first block revealed that children with ASD were just as likely as their TD controls to use the prosodic cue to inform ambiguous sentence interpretation. In the second block, however, children with ASD showed an inability to switch their offline actions to the new interpretation of the sentence when the prosody changed. On the surface, this failure to revise might indicate that children with ASD cannot reliably use prosodic cues to guide pragmatic interpretation; eye-gaze data from the second block, however, revealed that children with ASD did show anticipatory eye movements to the correct interpretations of the sentences in early time windows. This inconsistency between the offline actions and online eye-movements of children with ASD highlights a crucial clarification of the nature of the differences between the TD and ASD populations in the study: while both groups form online expectations about the ambiguous sentences, children with ASD may not able to flexibly shift their interpretations. Eye-gaze paradigms are crucial for providing moment-to-moment interpretations of spoken language, which, in turn, provides greater insight into the true nature of these developmental language disorders.

SOLUTIONS AND RECOMMENDATIONS

Eye-tracking technologies have shown us that in order to properly characterize language comprehension we need move beyond holistic offline responses and break to into the individual representational levels along the way and their interactions. The ability to pinpoint where along the processing stream language comprehension breaks down allows for a more nuanced understanding of how the system is structured and the sources of information that can aid or hinder final comprehension. One of the most striking conclusions that has been repeatedly supported by such research is that although the various mechanisms for language comprehension in children are remarkably similar to those of adults, there are distinct differences in when children are able to employ these mechanisms effectively and efficiently. On the whole, children's language processing system is slower; although children make predictions based on the incoming information to constrain their interpretation and help comprehension, the speed with which incoming information is processed lags behind that of adults. Further, across the different levels of processing, children's ability to use top-down cues such as referential context, and prosodic stress develops late and increases throughout the course of development. We would argue that for educational programs, consideration of these mechanistic differences between a child and adult listener is critical in order to identify difficulties at an earlier stage as well as creating interventions most likely to effect change.

Psycholinguistic research aims to test, better define, and even challenge current theories of language acquisition and development; theories which educational programming, for typically developing and developmentally delayed students alike, are grounded in. Although this chapter has highlighted some of the ways in which eye-tracking technology has improved our understanding of the complexities of language processing, much work is still needed to bridge the gap between research and its practical application. We encourage researchers, clinicians, and educators to use eye-tracking technologies to move beyond looking at the end states of language comprehension, understood versus not understood, to distinguish how each step of language processing is involved in order to reach a final sentence interpretation. In addition, although we discussed some recent work looking at language processing in children with atypical language development, much more research is needed in order to better understand and characterize the difficulties children with different language delays may face and the mechanisms that may allow for improvement through intervention.

FUTURE RESEARCH DIRECTIONS

Though we have briefly touched on some of the ways that eye-tracking technologies have allowed researchers to study the language deficits in children with developmental disorders, this area of research is still in its infancy. Eye-tracking methodologies continue to advance research on developmental disorders in two notable ways. First, as mentioned earlier in the chapter, eye-tracking measures have been shown to be valid indicators of online sentence comprehension in young children, and these measures are also reliably related to later language and cognitive outcomes. Continuing to research the ways in which early linguistic indicators can predict potential deficits later in development is imperative in order to establish earlier markers for children as risk for developmental delays. Second, beyond providing early indicators, eye-tracking research is also focusing on creating better-defined understandings of the developmental disorders themselves, uncovering important details in how different deficits may share similar mechanisms and how different mechanisms may conceal themselves in similar deficits. This research is critical for continued progress in creating effective treatments for children with developmental delays.

CONCLUSION

This chapter has highlighted some of the ways in which eye-tracking technology has been used to study spoken language comprehension in adults and children with and without developmental delays. The incredible speed and ease with which adult listeners make sense of the incoming speech stream often masks the complicated steps involved in the process. Eye-tracking has allowed researchers to break down the process into its component parts, working toward a better understanding of how each level of the language comprehension system develops from childhood to adulthood. Progress in psycholinguistic research over the past twenty years has revealed that though there are striking qualitative similarities between the child and adult language processing mechanisms, there are also differences in the types of information children can and cannot use in order to guide their construction of sentence comprehension. Eye-tracking technology is sure to play an important role in psycholinguistic research as investigators continue to focus on understanding when and how language comprehension mechanisms develop.

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 and Language*, *38*(4), 419–439. doi:10.1006/jmla.1997.2558

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: American Psychiatric Association.

Andreu, L., Sanz-Torrent, M., & Guardia-Olmos, J. (2012). Auditory word recognition of nouns and verbs in children with Specific Language Impairment (SLI). *Journal of Communication Disorders*, *45*(1), 20–34. doi:10.1016/j.jcomdis.2011.09.003 PMID:22055614

Arnold, J. E. (2008). The BACON not the bacon: How children and adults understand accented and unaccented noun phrases. *Cognition*, *108*(1), 69–99. doi:10.1016/j.cognition.2008.01.001 PMID:18358460

Bishop, D. (1997). *Uncommon understanding*. London: Psychology Press.

Brock, J., Norbury, C., Einav, S., & Nation, K. (2008). Do individuals with autism process words in context? Evidence from language-mediated eye-movements. *Cognition*, *108*(3), 896–904. doi:10.1016/j.cognition.2008.06.007 PMID:18692181

Brown, G. (1983). Prosodic structure and the given/new distinction. In A. Cutler & R. Ladd (Eds.), *Prosody: Models and measurements* (pp. 67–77). New York: Springer. doi:10.1007/978-3-642-69103-4_6

Choi, Y., & Trueswell, J. C. (2006). Children's (in)ability to recover from garden paths in a verb-final language: Evidence for developing control in sentence processing. *Journal of Experimental Child Psychology*, *106*(1), 41–61. doi:10.1016/j.jecp.2010.01.003 PMID:20163806

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology*, *6*(1), 84–107. doi:10.1016/0010-0285(74)90005-X

Dahan, D., Magnuson, J. S., & Tanenhaus, M. K. (2001). Time course of frequency effects in spoken-word recognition: Evidence from eye movements. *Cognitive Psychology*, *42*(4), 317–367. doi:10.1006/cogp.2001.0750 PMID:11368527

Dahan, D., Magnuson, J. S., Tanenhaus, M. K., & Hogan, E. M. (2001). Subcategorical mismatches and the time course of lexical access: Evidence for lexical competition. *Language and Cognitive Processes*, *16*(5-6), 507–534. doi:10.1080/01690960143000074

Dahan, D., Tanenhaus, M. K., & Chambers, C. G. (2002). Accent and reference resolution in spoken-language comprehension. *Journal of Memory and Language*, *47*(2), 292–314. doi:10.1016/S0749-596X(02)00001-3

Diehl, J., Friedberg, C., Paul, R., & Snedeker, J. (in press). The use of prosody during syntactic processing in children and adolescents with autism spectrum disorders. *Development and Psychopathology*. DOI: 10.1017/S0954579414000741

Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the second year. *Developmental Psychology*, *42*(1), 98–116. doi:10.1037/0012-1649.42.1.98 PMID:16420121

Fernald, A., Pinto, J. P., Swingley, D., Weinberg, A., & McRoberts, G. W. (1998). Rapid gains in speed of verbal processing by infants in the second year. *Psychological Science*, *9*(3), 72–75. doi:10.1111/1467-9280.00044

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), 1003–1015. doi:10.1111/1467-8624.00331 PMID:11480931

Frith, U. (1989). *Autism: Explaining the enigma*. Oxford, UK: Blackwell.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), Syntax and Semantics:*Speech Acts* (vol. 3, pp. 41–58). New York, NY: Academic Press.

Grodner, D. J., Klein, N. M., Carbary, K. M., & Tanenhaus, M. K. (2010). "Some," and possibly all, scalar inferences are not delayed: Evidence for immediate pragmatic enrichment. *Cognition*, *116*(1), 42–55. doi:10.1016/j.cognition.2010.03.014 PMID:20394919

Huang, Y. T., & Snedeker, J. (2009). Online interpretation of scalar quantifiers: Insight into the semantics-pragmatics interface. *Cognitive Psychology*, *58*(3), 376–415. doi:10.1016/j.cogpsych.2008.09.001 PMID:18976987

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*(6), 597–626. doi:10.1023/A:1026468209238 PMID:11196065

Ito, K., Bibyk, S. A., Wagner, L., & Speer, S. R. (2014). Interpretation of contrastive pitch accent in six-to eleven-year-old English-speaking children (and adults). *Journal of Child Language*, *41*(01), 84–110. doi:10.1017/S0305000912000554 PMID:23253142

Ito, K., Jincho, N., Minai, U., Yamane, N., & Mazuka, R. (2012). Intonation facilitates contrast resolution: Evidence from Japanese adults and 6-year olds. *Journal of Memory and Language*, *66*(1), 265–284. doi:10.1016/j.jml.2011.09.002

Ito, K., & Speer, S. R. (2008). Anticipatory effect of intonation: Eye movements during instructed visual search. *Journal of Memory and Language*, *58*(2), 541–573. doi:10.1016/j.jml.2007.06.013 PMID:19190719

Joseph, R. M., Tager-Flusberg, H., & Lord, C. (2002). Cognitive profiles and social- communicative functioning in children with autism spectrum disorder. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *43*(6), 807–821. doi:10.1111/1469-7610.00092 PMID:12236615

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*(04), 855–876. doi:10.1017/S0305000905007051 PMID:16429714

Kidd, E., Stewart, A. J., & Serratrice, L. (2011). Children do not overcome lexical biases where adults do: The role of the referential scene in garden-path recovery. *Journal of Child Language*, *38*(01), 222–234. doi:10.1017/S0305000909990316 PMID:20196901

Kjelgaard, M., & Tager-Flusberg, H. (2001). An investigation of language impairment in autism: Implications for genetic subgroups. *Language and Cognitive Processes*, *16*(2-3), 287–308. doi:10.1080/01690960042000058 PMID:16703115

Liversedge, S., & Findlay, J. (2001). Saccadic eye movements and cognition. *Trends in Cognitive Sciences*, *4*(1), 6–14. doi:10.1016/S1364-6613(99)01418-7 PMID:10637617

Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science*, *11*(3), F9–F16. doi:10.1111/j.1467-7687.2008.00671.x PMID:18466367

Marslen-Wilson, W., & Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology*, *15*(3), 576–585.

McMurray, B., Samelson, V. M., Lee, S. H., & Tomblin, B. (2010). Individual differences in online spoken word recognition: Implications for SLI. *Cognitive Psychology*, *60*(1), 1–39. doi:10.1016/j.cogpsych.2009.06.003 PMID:19836014

Minshew, N. J., Goldstein, G., & Siegel, D. (1997). Neuropsychologic functioning in autism: Profile of a complex information processing disorder. *Journal of the International Neuropsychological Society*, *3*, 303–316. PMID:9260440

Nation, K. (2008). Using eye-movements to investigate developmental language disorders. In C. F. Norbury, J. B. Tomblin, & D. V. M. Bishop (Eds.), *Understanding developmental language disorders* (pp. 39–52). New York, NY: Psychology Press.

Snedeker, J. (2009). Sentence Processing. In E. Bavin (Ed.), *The handbook of child language* (pp. 331–338). Cambridge University Press. doi:10.1017/CBO9780511576164.018

Snedeker, J., & Trueswell, J. C. (2003). Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language*, *48*(1), 103–130. doi:10.1016/S0749-596X(02)00519-3

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. doi:10.1016/j.cogpsych.2004.03.001 PMID:15342261

Snedeker, J., Worek, A., & Shafto, C. (2009). The role of lexical bias and global plausibility in children's online parsing: A developmental shift from bottom-up to top-down cues. Paper presented at the Boston University Conference on Language Development, Boston, MA.

Snedeker, J., & Yuan, S. (2008). Effects of prosodic and lexical constraints on parsing in young children (and adults). *Journal of Memory and Language*, *58*(2), 574–608. doi:10.1016/j.jml.2007.08.001 PMID:19190721

Tager-Flusberg, H., Paul, R., & Lord, C. (2005). Language and communication in autism. In F. R. Volkmar, R. Paul, A. Klin, & D. J. Cohen (Eds.), *Handbook of autism and pervasive developmental disorders: Diagnosis, development, neurobiology, and behavior* (3rd ed.; pp. 335–364). Hoboken, NJ: John Wiley & Sons, Inc.

Tanenhaus, M. K., Magnuson, J. S., Dahan, D., & Chambers, C. (2000). Eye movements and lexical access in spoken-language comprehension: Evaluating a linking hypothesis between fixations and linguistic processing. *Journal of Psycholinguistic Research*, *29*(6), 557–580. doi:10.1023/A:1026464108329 PMID:11196063

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*(5217), 1632–1634. doi:10.1126/science.7777863 PMID:7777863

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech and Hearing Research*, *40*(6), 1245–1260. doi:10.1044/jslhr.4006.1245 PMID:9430746

Trueswell, J. C. (2008). Using eye movements as a developmental measure within psycholinguistics. In I. A. Sekerina, E. M. Fernández, & H. Clahsen (Eds.), *Developmental psycholinguistics: On-line methods in children's language processing* (pp. 71–96). Philadelphia, PA: John Benjamins Publishing Company. doi:10.1075/lald.44.05tru

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*(2), 89–134. doi:10.1016/S0010-0277(99)00032-3 PMID:10580160

Weighall, A. R. (2008). The kindergarten path effect revisited: Children's use of context in processing structural ambiguities. *Journal of Experimental Child Psychology*, *99*(2), 75–95. doi:10.1016/j.jecp.2007.10.004 PMID:18070628

Young, E. C., Diehl, J. J., Morris, D., Hyman, S. L., & Bennetto, L. (2005). The use of two language tests to identify pragmatic language problems in children with autism spectrum disorders. *Language, Speech, and Hearing Services in Schools*, *36*(1), 62–72. doi:10.1044/0161-1461(2005/006) PMID:15801508

KEY TERMS AND DEFINITIONS

External Validity: The extent to which the results of a given study can be generalized to other situations and populations.

Felicitous: Well chosen. Appropriate, given the circumstances.

Linking Hypothesis: Series of assumptions that link a behavioral measure to a presumed cognitive function.

Metalinguistic: Conscious awareness of language and its function as a tool for communication.

Referential Context: Elements in the visual scene that relate to and provide information for a linguistic utterance.

Saccade: Quick eye-movements between fixation points.

Suprasegmental: Relating to features such as intonation, volume, or stress that co-occur with the phonetic features of an utterance.