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 to a large extent these assumptions are reasonable, 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 et al. 1998) and syntactic (e.g., Spivey et al. 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 resulting matrix of coordinates (triplets of pupil,

Figure 1. Head-mounted eyetracking

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.

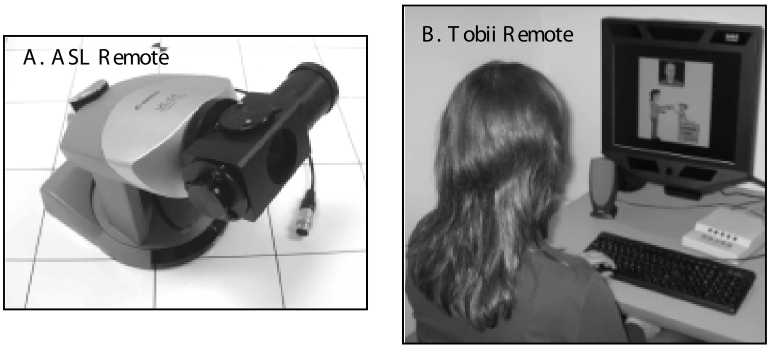

Figure 2. Remote eyetracking

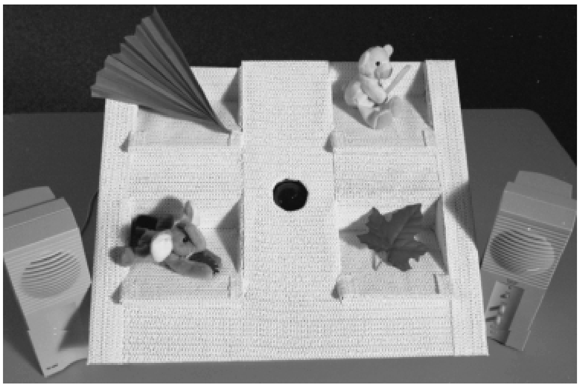

Figure 3. "Poor man's eyetracker"

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 eye-tracking 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 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 five 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*

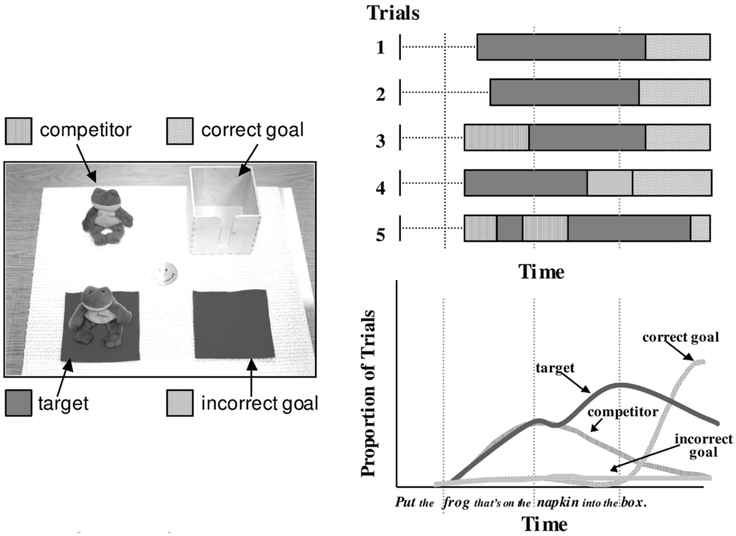

Figure 4. Calculating gaze proportions over time (modified from Tanenhaus & Trueswell 2005)

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 what are happening in the eye movement record.

The actual data from this condition in the Trueswell et al. study is 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; first, looks

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.

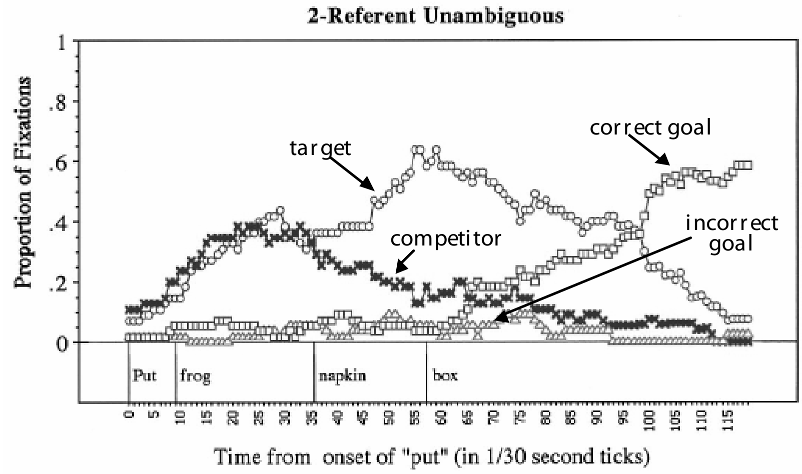

Figure 5. Children's (5 year olds') proportion of looks over time to potential referent objects in response to *Put the frog that's on the napkin into the box*. From Trueswell, Sekerina, Hill and Logrip (1999). Copyright 1999, Elsevier Press.

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 et al 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 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 (Fernald & Hurtado 2006; see also Fernald 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 '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 straight-forward 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 tasks. Without knowledge of these facts, one will not be able to adequately inter-

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.

pret 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 four 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 straight-forward 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 instead 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 one to four 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 above, 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 simply 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. 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 simplicity; 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

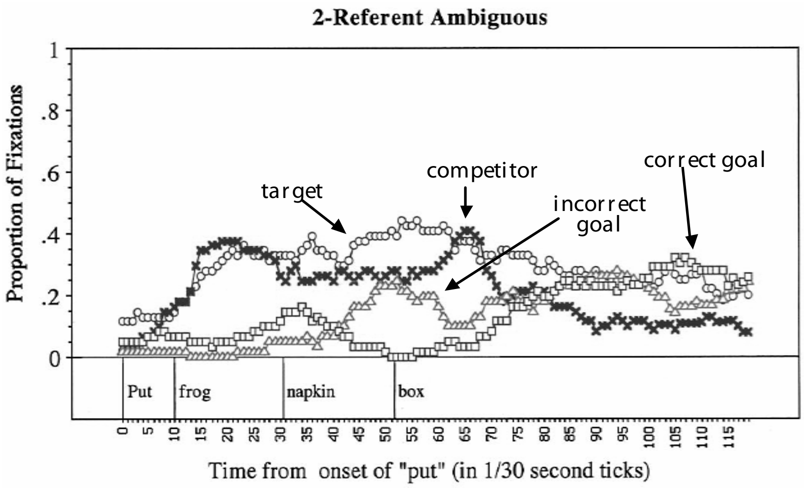

Figure 6. Children's (5 year olds') proportion of looks over time to potential referent objects in response to *Put the frog on the napkin into the box*. From Trueswell, Sekerina, Hill and Logrip (1999). Copyright 1999, Elsevier Press.

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 (in submitted) 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 2007) 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 twelve years of age and younger show difficulty reinterpreting situations and inhibiting pre-potent responses to stimuli (e.g., Bialystok 2001; Bunge, Dudukovic, Thomason, Vaidya, & Gabrieli 2002). For instance, children under the age of six years show difficulty discovering the alternative interpretation of ambiguous figures (Bialystok & Shapero 2005). Similarly, children in this age range show difficulties over-riding 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 for these and related experimental findings see Davidson, Amso, Anderson, & Diamond 2006, and references therein.)

Put another way, children are surprisingly 'cognitively impulsive.' Automatic and/or highly learned responses to stimuli are often difficult for a child to rescind and revise. This behavioral pattern over development follows 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 five 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 five 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 five 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) explores 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 five 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. In particular, 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 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.

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.* 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.

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.

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.

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. (submitted). Eye movements during spoken-word recognition in Russian children. Manuscript submitted for publication.
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. (2007). Children's representation of verbs: Evidence from priming during online sentence comprehension. Manuscript submitted for publication.

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.