Linking Eye Movements to Sentence Comprehension in Reading and Listening

Julie E. Boland
University of Michigan

Eye tracking paradigms in both written and spoken modalities are the state of the art for online behavioral investigations of language comprehension. But it is almost a misnomer to refer to the two types of paradigms by the same "eye-tracking" label, because they are quite different. Reading paradigms gauge local processing difficulty by measuring the participant's gaze on the very material that he or she is trying to comprehend. The critical sentence regions are determined spatially, and gaze is measured in terms of the time spent looking within a region of interest, the likelihood of a regressive eye movement out of the region, and so forth. In contrast, listening paradigms gauge how rapidly successful comprehension occurs by measuring how quickly people look, or how likely people are to look, at objects referenced by the linguistic material.

This chapter summarizes some of the contributions of each paradigm, focusing on the linking assumptions between eye movement data and sentence comprehension processes. I will also discuss some limitations that currently plague each paradigm and make a few suggestions for how we might get to the next level of investigation using these paradigms. The first half of the chapter focuses on reading, and the second half on listening. It will quickly become clear that the same linking assumptions won't serve us in both reading and listening paradigms. Consider Figure 1. On one side of this figure are some of the cognitive processes that might be under investigation. On the other side are the behaviors that we can measure in an eye movement paradigm. Intuitively, the linkage between the two sides is straightforward in reading paradigms: Reading comprehension involves visual attention, and visual attention requires fixation. However it is less clear,

on the face of it, why eye movements should be linked to language comprehension in a listening paradigm.

Insert Figure 1 about here

EYE MOVEMENTS AND READING

Eye movement data from reading have been very influential for evaluating theories of human sentence processing. This is because the eye movement record provides an on-line measure of processing difficulty with high temporal resolution, without relying on any secondary task to produce the dependent measures. Furthermore, we know a lot about factors that influence the planning and execution of saccades, and the duration of fixations, because eye movements have been carefully studied for several decades within the domains of vision, motor control, and language processing (see Rayner, 1998, for a recent review). For example, researchers have studied how our eyes move over meaningless strings, and then examined how much of the variability in fixation duration and saccade landing site is linked to linguistic factors in actual text. In fact, we have a pretty clear understanding of the amount of visual information processed during a fixation on text. During reading, factors such as a word's frequency, length, predictability, and ease of integration into the sentence influence how long it takes to access the lexical entry for that word and to incorporate the new lexical information into the structural and conceptual representations the reader is constructing for the sentence (e.g., Pollatsek & Rayner, 1990; Rayner et al., 1996). These same factors also influence whether the eyes fixate on a word and, if so, how long the fixation is maintained (Just & Carpenter, 1980; Rayner et al., 1998; Rayner et al., 1989; Reichle et al., 1998). In contrast, we don't understand how the different dependent measures that are commonly used (e.g., first fixation time, probability of a regression) are linked to specific cognitive events.

[insert Table with dependent measures and definitions? Is this covered in Martin's chapter? If so, I will simply refer readers there.]

Multiple Dependent Measures: Converging and Diverging Evidence

It is undoubtedly an advantage that we can carve the eye movement record up in different ways. Most researchers analyze three or four different dependent measures in hopes of getting a complete view of the cognitive processes involved in sentence comprehension. Oftentimes, they find the same pattern in all the dependent measures, which makes for a coherent set of results. However, I will argue that we can learn more about the linking assumptions between eye movement data and the cognitive events from sentence comprehension by studying experiments in which the dependent measures diverge.

An example of converging evidence is the classic paper by Frazier and Rayner (1982). Participants read garden path sentences like those in (1). It is well known that readers and listeners develop structural (syntactic) and interpretive (semantic) representations of sentences incrementally, as they read or hear each word. However, the structural position of *a mile* is temporarily ambiguous; it can either be the direct object of *jogs* as in (1a), or the subject of an embedded clause, as in (1b).[1]

(1) a. *Since Jay always jogs a mile <u>this seems</u> like a short distance to him.*

b. *Since Jay always jogs a mile <u>seems</u> like a short distance to him.*

The experimenters expected readers to adopt the direct object structure and to experience processing difficulty in (1b) at the point of disambiguation, which is underlined. As predicted, *seems* was fixated longer and regressive eye movements occurred more frequently in (1b) than in (1a). Some trials exhibited the effect in one dependent variable, some trials the other, and some trials exhibited the effect in both dependent variables. The combination of long fixations and regressive eye movements at the point of disambiguation has become the hallmark of a garden path in eye movement studies of

reading and such effects can be found even when the reader is unaware of any processing difficulty.

Much more recently, Traxler et al. (2002) found that the difficulty of reanalysis after a garden path in object relative sentences was mitigated by animacy. This pattern was seen as a trend in the first pass data, and was reliable in their other three dependent measures (quasi-first pass, the likelihood of a regression, and total reading time). Findings like these, of which there are many, suggest that processing load increases are equally likely to be reflected in longer fixations, secondary fixations in the difficult region, and regressions to earlier segments.[2] The linking assumption is then fairly simple: A single type of cognitive event can result in various behavioral outcomes.

When an effect observed in the total reading time is also present in the first fixation data, it demonstrates that the relevant constraint affects initial processing of a given region. For example, if an animacy effect is present in the first fixation on a word, one can conclude that animacy affected processing of that word during or soon after word recognition. However, the following questions may be more easily answered when the various dependent measures do not all exhibit the same pattern.

- Do some types of constraints influence syntactic ambiguity resolution, but not the initial generation of structural alternatives?
- Which constraints determine the structure(s) that are initially accessed/constructed and how easy it is to do so?
- Are phrase structure, morpho-syntactic, and semantic operations sequentially ordered or simultaneous?

It is fairly common in reading studies to contrast early (first pass) effects with later (second pass) effects. In doing so, one hopes to use the multiple dependent variables

offered by the eye movement record to distinguish early structure-building processes from later processes that make use of those structures. Importantly, different researchers have taken this contrast to reflect different cognitive events, depending upon their theoretical assumptions. For example, some researchers have argued that detailed lexical information is not part of the initial structure building process (e.g., Mitchell, 1987) while others have argued that detailed lexical information forms the basis of the initial structure building process (e.g., MacDonald et al., 1994). In my own research, I have argued that the architecture of the sentence comprehension systems is restricted, such that lexical and syntactic constraints influence the initial construction of syntactic alternatives, as each new word is integrated into the developing sentence structure. At the same time, I have maintained that constraints from higher levels of representation can influence syntactic ambiguity resolution when there are multiple structural alternatives generated. The first of these two architectural claims was tested and supported by Boland and Blodgett (2001).

Boland and Blodgett (2001) embedded unambiguous target sentences like (2a) and (2b) in simple stories. The sentences contain a noun/verb homograph ("duck"), but its syntactic category is disambiguated by the preceding pronoun.

(2) Example target sentences from Boland & Blodgett, 2001. Bars separate regions used for analysis.

a) Noun target. *She* | *saw his* | *duck and* | *chickens near* | *the barn.*

b) Verb target. *She* | *saw him* | *duck and* | *stumble near* | *the barn.*

The examples in (2) are taken from a story about a girl visiting a boy on a farm. The sentence prior to the target sentence supported either a noun or verb meaning of the homograph: "Kate watched everything that Jimmy did." Or "Kate looked at all of Jimmy's pets." The type of context sentence and the type of target sentence were crossed to create four levels of discourse congruency. In addition, the relative frequency of the

noun and verb forms of the homograph were varied continuously to create a lexical frequency variable. Following the architectural assumptions outlined above, we predicted that the more frequent the appropriate form of the homograph, the easier it would be to integrate the homograph with the developing structural representation. In contrast, we expected that discourse congruency would not affect ease of structure building in this unambiguous syntactic context. Using the linking assumption in (3), this led to the empirical prediction that only lexical frequency effects would be seen in the first pass reading times of these unambiguous target sentences.

(3) Linking assumption: First fixation duration reflects ease of structural integration, but not pragmatic/discourse integration.

As predicted, there was a lexical frequency effect in the first fixation on the region "duck and"—the earliest possible location. In contrast, no discourse congruency effect was seen in the first pass reading times. There was a discourse congruency effect in the likelihood of regressive eye movements, but this effect was downstream, in the region after the lexical frequency effect. There was also a very robust effect of discourse congruency in the second pass reading times, distributed throughout most of the sentence. Note that in this contrast between early and late effects, there is a difference in both the dependent variable that exhibited the effects and in the sentence region where the effect was found.

The early vs. late contrast in Boland and Blodgett (2001) has theoretical implications, because it suggests that lexical frequency influences the initial generation of structure, but discourse congruency does not. The later effects of discourse congruency are assumed to reflect anomaly detection after the sentence structure has been determined. An important part of the argument is that discourse congruency DOES have immediate effects on structural ambiguity resolution. In locally ambiguous sentences like "She saw her duck…" Boland (1997) found immediate effects of discourse congruency at the point

of disambiguation (which was either the word following the ambiguous pronoun or downstream of the noun/verb homograph). This demonstrates that discourse congruency does resolve a syntactic ambiguity when a congruent and an incongruent structure are both syntactically possible, but that discourse congruency can't determine what structures are possible, nor does it influence the initial structure generating process.

Anomaly Detection

In Boland and Blodgett (2001), the discourse anomaly effects were observed relatively late in the eye movement record. Under what circumstances should anomaly effects arise late? Would local structural anomalies reveal themselves earlier than discourse anomalies? To fully answer these questions would require a detailed model of the mapping between the dependent measures and the underlying cognitive processes. While no such model exists, there is some relevant evidence that will be reviewed in this section.

It should be noted that eye movement research has not distinguished itself as much as some other methodologies in the area of anomaly detection. In particular, event-related potential (ERP) research has been quite promising in distinguishing anomaly detection processes in terms of the linguistic level of analysis at which the anomaly occurs (e.g., Ainsworth-Darnell et al., 1998; Friederici & Frisch, 2000; Gunter et al., 1997; Hagoort et al., 1993; Osterhout & Nicol, 1999; Rosler et al., 1993). [Can I refer the reader to Lee's or Jos's chapters for ERP anomaly detection summary?] Most readers are probably familiar with the N400 component as an index of semantic anomaly/predictability and the P600 (along with some earlier components) as an indicator of syntactic anomaly. In fact, Friederici (1995) outlined a detailed processing architecture in which phrase structure construction (based on major syntactic category, as in Frazier's (1978, 1987) garden path theory) precedes morpho-syntactic and lexical-semantic processing. Her proposed

architecture is motivated by differences in both the scalp location and the latency of ERP anomaly detection effects. Given the latency differences observed in the ERP record, should we not see similar distinctions in the eye movement record, where we have more detailed temporal information?

In fact, most eye movement studies have focused on garden path sentences rather than the globally anomalous sentences typically used in ERP research. A garden path occurs when a temporarily ambiguous region of a sentence is misanalyzed. The initial analysis becomes anomalous when disambiguating words are encountered later in the sentence. However, readers can often eliminate the anomaly by restructuring the ambiguous portion of the sentence. As noted above, garden path effects are characterized by increased looking time in the disambiguating region and higher probabilities of regressions out of the disambiguating region (e.g., Frazier & Raynor, 1982). Because garden paths are generally taken as evidence that the language processing system pursues only one analysis of the ambiguous material, the linking assumption is that the cognitive phenomena of anomaly detection and syntactic reanalysis are reflected in the eye movement data as longer first pass reading times and an increased probability of regressive eye movements.

As noted by Ni et al. (1998; see also Fodor & Inoue, 1994, 2000), anomaly detection plays a central role in the garden path experience. The anomaly at issue may be a breakdown in sentence structure, as in Frazier and Rayner (1982), but sometimes it is a semantic violation, as in (4), from Rayner et al. (1983; see Ni et al., 1996, and Spivey-Knowlton & Sedivy, 1995, for more examples).

(4) *The spy saw the cop with a <u>revolver</u>, but the ….*

Logically, anomaly detection must precede reanalysis. Thus, under a serial analysis theory, anomaly detection must be linked to long initial fixations on the disambiguating

material. Such an assumption was made explicit by Frazier and Rayner (1982, p. 193): "When the reader was garden-pathed, there was an awareness at some level on the first fixation in the disambiguating region that something was wrong, as evidenced by a longer fixation duration." Because regressive eye movements often landed in the ambiguous region, Frazier and Rayner suggested that such regressions "indicate that subjects have detected an error in their initial analysis of the sentence and have identified the source of the error" (p. 203).

These particular links between eye movements and the underlying cognitive behavior have been widely accepted over the past two decades. However, there have been some cautionary notes and some conflicting conclusions regarding the interpretation of eye movement data. For example, Rayner et al. (1989) noted that regressions "could reflect not only the existence of an erroneous initial analysis, but also the relative plausibility of competing analyses, the syntactic differences between the initial and the revised analyses, and so on" (p. 38). In fact, Spivey and Tanenhaus (1998) attributed increased reading times to competition between parallel structures--not to anomaly detection and reanalysis. Their constraint-based lexicalist account allows such an attribution because, in contrast to Frazier and Rayner (1982), they assume that syntactic alternatives are postulated and evaluated in parallel. In short, despite some commonly held assumptions, there are many open research questions concerning the relationships between the dependent measures in eye movement paradigms and the cognitive processes underlying sentence processing.

If noticing something anomalous is the initial component of processing difficulty in garden path sentences, then one would expect to see anomaly effects in the first pass fixations on the words that induce the anomaly in both garden path sentences and globally anomalous sentences like those typically used in ERP experiments. However, a review of the small eye movement literature on unambiguous, globally anomalous

sentences suggests that this is not always true. Table 1 provides a summary of anomaly detection effects from the eye movement literature. Only local first pass measures are considered here. The shaded areas in Table 1 represent the types of anomalies that were investigated in each study. Unfortunately, there have not been many papers that investigate several types of anomalies within a single experiment. This makes it difficult to determine (in analogy to the ERP experiments) whether there is an "anomaly detection pattern" of eye movements, and whether the gaze response to an anomaly differs depending upon linguistic level of anomaly.

Insert Table 1 about here

Boland and Blodgett (2002) examined the broadest range of anomaly types. They used two sets of critical stimuli, an "argument structure set" and an "agreement set." The first set of stimuli contained verb argument structure violations on the indirect object. Examples are given in (5) and (6). Anomalous words are starred and doubly anomalous words receive two stars. The critical word at which the potential anomaly is apparent is underlined. The anomalous indirect object was either of the wrong phrasal category—a noun phrase (NP) instead of a prepositional phrase (PP) as in (5b)—or had the wrong semantic properties (*signs*), or both. The anomalies in the agreement stimuli were either a syntactic subject-verb agreement violation as in (6b) or a semantic violation in which the sentential subject (*snake/s*) was not a suitable agent for the verb. Thus, in both stimulus sets, the semantically anomalous words were inconsistent with the thematic role that the verb assigned to one of its arguments. However, the syntactic violations were quite different. The syntactic anomalies in the agreement stimuli were morpho-syntactic agreement violations similar to those used by Ni et al. (1998), Braze et al. (in press), and Pearlmutter et al. (1999). In the argument structure stimuli, the violations reflected a conflict between the verb's argument structure and the phrasal category of the indirect

object.

(5) a. *Kim recommended Shakespeare to <u>everyone/*signs</u> after she saw Hamlet.*

b. *Kim recommended Shakespeare *<u>everyone/**signs</u> after she saw Hamlet.*

(6) a. *The canary/*snake in the large cage <u>sings</u> beautifully*.

b. *The *canaries/**snakes in the large cage <u>sings</u> beautifully*.

As noted above, there is some evidence from the ERP literature that phrasal category violations are recognized more rapidly (e.g., Friederici et al., 1993), but there have been no previous attempts in the eye movement literature to distinguish between morpho-syntactic and phrasal category violations. Calling to mind the Friederici distinction[3], Boland and Blodgett (2002) found that syntactic congruency affected first pass reading time when the manipulation involved a phrasal category error, but not when it involved a morphological feature error. Both types of syntactic congruency influenced the likelihood of a first pass regressive eye movement. Semantic congruency did not influence first pass reading time, but it did affect the regression path duration[4].

Ni et al. (1998) investigated syntactic and semantic anomalies in unambiguous sentences like those in (7). The critical word is underlined, and ungrammatical sentences are starred. Both the syntactic and the semantic anomalies (illustrated in (7a) and (7b), respectively) led to more regressive eye movements compared to the control sentences (7c). The likelihood of an immediate regression, however, was higher for syntactic anomalies than for semantic anomalies. In contrast, only the semantic anomaly induced longer first pass fixations, and only after (rather than during) the anomalous region. Ni et al. interpreted these results to suggest qualitative differences in the cognitive response to the two types of anomalies—differences that are directly reflected in the eye movement

patterns.

(7) a. **It seems that the cats won't usually eating the food . . .*

b. **It seems that the cats won't usually bake the food . . .*

c. *It seems that the cats won't usually eat the food . . .*

More recently, Braze et al. (in press) focused on regressive eye movements in comparing morpho-syntactic and semantic anomalies. They found that syntactic anomalies elicited an immediate peak in the percentage of regressive eye movements, while semantic anomalies led to a gradual rise in regressions that peaked at the end of the sentence. Braze et al. also reported increased reading times just after the semantically anomalous word.

Pearlmutter et al. (1999, Experiment 2) examined subject-verb agreement errors in sentences like (8). As in Ni et al. (1998) and Braze et al. (in press), the eye movement pattern was dominated by regressions. However, Pearlmutter et al. did find longer first pass fixations for some conditions on a subset of trials. The effect was not localized to the anomalous word alone (*were*), but rather became apparent when fixations on the anomalous word and the following word were summed. Furthermore, the effect was only observed after excluding the trials on which there was a regression out of that region (about 14% of trials), and excluding eight participants who either always or never made regressions out of that region.

(8) a. *The key to the cabinet was/*were rusty…*

b. *The key to the cabinets was/*were rusty…*

Deutsch and Bentin (2001) examined subject-verb gender agreement in Hebrew. In contrast to the English morpho-syntactic anomaly results, they found a first pass anomaly effect for the marked (plural verb) form (9a). Only second pass effects were observed for the unmarked (singular) form (9b). The authors also reported an ERP version of the

experiment. They found a negative component 80 – 250 ms after incongruent verbs, regardless of markedness, and a P600 for marked incongruent verbs only.

(9) a. *I enjoyed seeing how <u>the actors</u> (*hasaxkani**ot**/hasaxkan**im**) were enchanting (maksim**im**)…*

b. *The woman saw that the <u>boy/girl</u> (*hayeled/hayalda) had fallen (nepal)…*

Frisson and Pickering (1999) reported some early and some late semantic anomaly effects in their investigation of metonymic expressions. Familiar metonomy ("Americans protested during <u>Vietnam</u>.") is handled easily; the authors report that readers were able to coerce the place (Vietnam) into an event (the Vietnam War) with no increase in processing difficulty over a literal control ("Americans hitchhiked around Vietnam."). However, unfamiliar metonymic expressions seem to be treated as semantic anomalies, as in (10a), *Finland* condition. As in the Boland and Blodgett experiment, there was no evidence of an anomaly in the first pass reading times for items like (10a). In contrast, Frisson and Pickering reported local first pass reading time effects for semantic anomalies like those in (10b). The crucial difference between (10a) and (10b) is unclear. The authors suggest that "Finland" in (10a) initiates a search for relevant events, while there is a clear feature clash or selectional restriction in (10b). However the details of such an account need to be carefully worked out. Why should unfamiliar "place for institution" metonomy violate a selectional restriction if unfamiliar "place for event" metonomy does not?

(10) a. *A lot of Americans protested during *<u>Finland/Vietnam</u>…*

*b. The blasphemous woman had to answer to the *<u>stadium/convent</u>…*

What generalizations can we make about these anomaly effects in unambiguous sentences? Can we generate any predictions about the types of anomalies that will generate local first pass effects? In the one unpublished study that examined phrasal

category violations, eye movement evidence of the violations emerged early, as most sentence processing theories would predict (Boland & Blodgett, 2002). That is, most theories maintain that lexically sensitive structure generation is an essential component of early sentence comprehension processes. However, not all syntactic violations led to immediate effects. There was much more data on morpho-syntactic errors, such as agreement violations. The English studies consistently find morpho-syntactic effects in the probability of a regression, while they don't find increased first pass reading times on the anomalous word (Boland & Blodgett; Ni et al., 1998; Braze et al., in press; Pearlmutter et al., 1999). In contrast, a Hebrew study did find first pass reading time effects for some of the anomalous conditions (Deutsch & Bentin, 2001). The results regarding semantic/pragmatic violations are also mixed. Several studies found no local first pass effects (Boland & Blodgett; Braze et al.; Ni et al.), but Frisson and Pickering (1999) did find first pass reading time effects for certain semantic violations. (These semantic violation studies are all in English.)

Box 1. Pushing the Linking Assumptions to the Next Level
Working hypothesis: The eyes don't leave a word until it has been structurally integrated. Thus, constraints that control structure building affect first pass time.

- Syntactic category, subcategorization
- Lexical frequency
- Morphological agreement in richly case-marked languages (like Hebrew, but not English)
- Semantics, when it determines which structure is constructed (?)

Box 1 offers a strong linking assumption as a working hypothesis: *The eyes don't leave a word until it has been structurally integrated. Thus, constraints that control structure building affect first pass reading time*. Uncontroversially, major syntactic category is a constraint that influences structure building. There is now considerable evidence that subcategorization and lexical frequency also influence structure building directly (see

Tanenhaus & Trueswell, 1995). Agreement is not a structure-determining constraint in English: when agreement features do not match, the structure is ungrammatical, but we know what the structure is. However, in languages with freer word orders, like Hebrew, the agreement features are used to determine the structure. Semantic constraints sometimes determine local structure, but often do not. Note that the relevant constraints summarized in Box 1 are quite different from the traditional "first pass" assumptions by Frazier and colleagues in garden path literature (e.g., Frazier, 1978; 1987). In Frazier's garden path model, neither detailed lexical information (subcategorization, frequency information) nor semantics could influence the initial rule-based parse.

Even if it turns out to be right, the linking assumption in Box 1 leaves much to be worked out. For example, what is the relationship between regressive eye movements out of a region and long/repeated fixations within that region? Why should first pass reading time be linked to initial structure generation rather than complete understanding? The proposed linking assumption implies that only structure determining factors will increase first pass time, but there are obvious exceptions to that generalization, such as the discourse level effects reported by O'Brien et al. (1988) and Garrod et al. (1994). If the assumption in Box 1 is to be maintained, one must identify the conditions under which these effects arise early.

The Promise and Limitations of Reading Paradigms

Most of eye movement research in reading has focused on garden path sentences, leading to many insights about syntactic ambiguity resolution. However, because garden paths involve both anomaly detection and reanalysis, it is difficult to establish the mapping between cognitive events and eye movement behaviors. In this chapter, I have focused on eye movement patterns over unambiguous sentences as a means to contrast constraints on initial syntactic generation with constraints on other processes, such as syntactic

ambiguity resolution, morphological agreement (in languages like English), and semantic/pragmatic felicity. I have argued that first fixation and first pass reading times in unambiguous sentences may provide a relatively pure index of syntactic generation. Thus, first pass duration is influenced by lexical frequency, but not discourse congruency (Boland & Blodgett, 2001), consistent with an architecture in which structure generation is largely lexical, with discourse constraints operating upon the lexically generated structures. Likewise within anomaly detection paradigms, first pass duration is influenced by (sub)category and other structure-determining constraints.

Despite some measure of promise, it should be clear that there has been no definitive mapping between specific cognitive processes and eye behavior. Many questions remain unanswered. Do (specific) cognitive events trigger a saccade? Is there some cognitive event(s) that causes the eyes to regress rather than move forward? How much cognitive processing occurs in the parafovea? The answer to this last question is linked to another set of questions: To what degree are words recognized and structured in parallel? If there is some degree of parallelism, should we still think of the first fixation on a word as the earliest possible measure of processing for that word? Are non-fixations on a word cognitively meaningful?

In contrast to the ERP literature, there is no evidence that syntactic anomalies disrupt the eye movement record in one way and semantic anomalies disrupt it in another. We may always have to rely upon converging evidence from ERP or MEG to understand which levels of representation have registered an anomaly. Although eye tracking paradigms lack precision in addressing that kind of "how" question, they excel at answering the "when" questions, providing a detailed temporal record of local processing difficulty.

An inherent limitation is actually the reading itself. When relying upon reading as an

intermediary skill in order to measure language comprehension, one immediately limits one's population to that of skilled readers (omitting children, aphasics, etc.). Furthermore, many interesting questions about language comprehension can only be answered within listening paradigms and/or by "situating" language within a real world context. In the limit, one would like to study language comprehension in conversation, its original and most natural context. To address these kinds of concerns, spoken language paradigms have a great deal of appeal.

EYE MOVEMENTS & LISTENING

Over the last five to ten years, psycholinguists have been exploiting Cooper's (1974) finding that we tend to look at things as they are mentioned. Cooper used a passive listening task; he presented a set of pictures in a grid and found that listeners' looks to objects were time-locked to mention of those objects in a story. Some researchers make the link between mentioning an object and looking at it more explicit by asking listeners to move or point to objects, within a directed action paradigm. Regardless of the task (directed action vs. passive listening), eye movements in listening paradigms do not provide the multi-layered dependent measures found in reading paradigms, because the looks of interest are not overlaid onto some regular sequence of eye movements, as in reading. In listening paradigms, the dependent measures are usually limited to the probability and duration of a look to a relevant object within some temporal interval.

In reading research, a natural starting point was to study how the eyes move over meaningless strings, and then examine how the variability in fixations and saccades is associated with linguistic factors in actual text. No comparable approach has been taken in the listening literature, though one could—and perhaps ought to—in the struggle to outline the linking assumptions between visual attention and listening comprehension.

For now, the most we can do is look to the substantial literature on scene perception in order to ground our predictions about when and how eye movements should occur.

Research on scene perception has established that scenes are identified within the first fixation (Biederman et al., 1982; Boyce & Pollatsek, 1992; Hollingworth & Henderson, 1998). That is, basic global information about the scene is absorbed within a single fixation. The mean fixation duration in scene perception is about 330 ms, with a mode of 230 ms (Henderson & Hollingworth, 1999). The initial fixation patterns are quite similar across participants, but variance increases rapidly over time (Mannan et al., 1995). The locations of the initial fixation on a scene are based on visual, not semantic, features (Henderson & Hollingworth). However, the likelihood of refixation is based upon task-dependent informativeness (Henderson & Hollingworth). Finally, our working memory representation of scene is abstract/conceptual rather than strictly visual (Henderson & Hollingworth).

The scene perception literature suggests that the first second or so of looking at a scene is relatively uninteresting, at least within the range of instruction conditions typically used in scene perception studies.[5] Looks during this time are driven by visual parameters, with little variance in performance. In language comprehension experiments, the time interval of interest generally follows scene presentation by several seconds, so we should expect listeners to have already scanned the scene and have an abstract representation of it. This mental representation allows listeners to look at maximally informative regions of the scene in response to linguistic input. In fact, the Cooper (1974) results suggest a simple linking assumption: *The probability of looking at an object increases when the object is mentioned.* This linking assumption has ecological validity in conversation, because we need to know properties of referents beyond their linguistic label. Directed action tasks further increase the probability of a look by using eye-hand coordination to mediate

comprehension and eye movements. That is, it is difficult to manipulate or point to an object without first looking at it.

In short, listeners have reason to look at referents as they are mentioned in both passive listening and directed action tasks. This phenomenon has proven useful and informative in a variety of domains, as illustrated in Box 2. Research on spoken language comprehension has shown that eye fixations are time-locked to lexical access of isolated words (e.g. Allopenna et al., 1998), identification of referents for syntactically ambiguous phrases (e.g., Tanenhaus, et al., 1995; Novick & Trueswell, 2001), and pronoun resolution in discourse context (e.g., Arnold et al., 2000). Eye tracking experiments investigating language comprehension have been conducted on children (Trueswell et al., 1999) and in the context of conversation among adults (Brown-Schmidt et al., 2002).

Box 2. Important Contributions

- Cohort competition & frequency effects in lexical access (Allopenna et al., 1998; Dehan et al., 2001)
- Incremental reference assignment (Cooper, 1974; Eberhard et al., 1995)
- Visual referential context influences ambiguity resolution (Tanenhaus et al., 1995)
- Young children don't use visual context as efficiently in ambiguity resolution (Trueswell et al., 1999)
- Initial forays into conversational interaction (Brown-Schmidt et al., 2002)

Clearly, measuring eye movements in a listening paradigm can provide a sensitive index of referential success or ambiguity resolution once linguistic input has been comprehended. However, it is not clear that eye fixations are time-locked to developing syntactic and semantic representations in a manner that can be distinguished from the listener's ultimate conceptual representation of the linguistic input. Can this research paradigm be used to investigate how people develop linguistic representations of spoken

language input? Can we use it to unpack the cognitive operations in syntactic processing? Is it "on-line" enough to investigate parsing?

Anticipatory looks

Intriguingly, reference resolution is sometimes completed prior to actual mention of the referent. For example, Sedivy et al. (1999, Cognition) found that when listeners were asked to *Pick up the tall glass…,* they often planned an eye movement to the glass during the adjective, prior to hearing "glass". Instead of waiting for bottom-up evidence of the target object, listeners used the current visual context (i.e., presence of a tall/short contrast set) to select the pragmatically appropriate referent. In a more recent study, Sussman et al. (2002) found that listeners made an eye movement to an appropriate instrument (a pencil) when hearing "Poke the dolphin" but not "Touch the dolphin." Thus, even though no instrument was mentioned, listeners used their knowledge about the two verbs to decide whether to manipulate the dolphin with their finger or a pencil in the real world environment. Listeners in the Sussman et al. study were also sensitive to contextual factors that altered verb meaning. For example, they looked at a potato peeler when asked to "Peel the potato", but not when asked to "Peel the banana."

Does argument structure implicitly introduce new entities into discourse?

Anticipatory looks such as these may provide clues to intermediate representations. For example, Sussman et al. (2002) concluded that hearing a verb provides access to its thematic grids and listeners then use context to select the appropriate grid. If the relevant thematic grid contained an instrument, eye movements were observed to an appropriate instrument in the environment. A related phenomenon has been reported in reading studies (Mauner et al.'s (1995) implicit agents; Carlson & Tanenhaus's (1988) open thematic roles). Just as in Sussman et al., it was argued that thematic role information from verb argument structure was accessed during word recognition. In the reading studies, the thematic roles guided comprehension even if an argument was not explicitly

mentioned. Together, these studies suggest that the thematic roles accessed during verb recognition can be used both to interpret current discourse entities and to introduce new entities into the discourse.

Of course, directed action tasks such as that used by Sedivy et al. (1999) and Sussman et al. (2002) give rise to the concern that listeners are strategically guessing the speakers intent, rather than allowing language processing to proceed automatically. Arguably, normal conversation involves a great deal of strategic guessing about the speaker's intent, so this is not necessarily a problem with the paradigm if the intent is to study the complete comprehension process. However, if there are some partially or fully automatized aspects of syntactic and semantic processing, the directed action paradigm is not ideal for studying the representations that results from those automatized processes alone. For example, one might question whether the recognition of "poke" **obligatorally** introduces an instrument into the discourse model. And importantly, are the discourse elements that can be introduced by the verb limited to members of its thematic grids? In other words, do the verb's arguments hold any kind of privileged status or are all related words/concepts accessed in the same way?

Encouragingly, work from Gerry Altmann's lab, as well as some of my own recent work, provide converging evidence for automatic activation of thematic role information from passive listening tasks. In an extremely interesting study, Altmann and Kamide (1999) recorded eye movements as people listened to a sentence like *The boy will move/eat the cake* and looked at a semi-realistic scene with a boy, a cake, and other moveable but not edible objects. Altmann and Kamide found faster looks to the cake following *eat* compared to *move*. In fact, participants often looked at *cake* in the *eat* condition prior to the onset of the noun. Altmann and Kamide concluded that the verb's thematic roles were used to pro-actively restrict the domain of subsequent reference.

Even in a passive listening task, it is difficult to identify the cause of the anticipatory fixations on the cake, because both linguistic and general world knowledge could have contributed to the effect. If it is solely the verb's argument structure that is driving eye movements, then listeners should not look at a bed upon hearing *The girl slept* because *bed* cannot be an argument of *slept*. However, beds are part of a prototypical sleeping event and are thus conceptually related to *sleep*. Furthermore, discussions about sleep often include mention of a bed, so linguistic co-occurrence frequency is high and the co-occurrence of sleeping and beds in participants' actual experience is likely to be extremely high. One might consider an account of Altmann and Kamide's (1999) effect that is akin to semantic priming--a conceptual, essentially intra-lexical, process. However, in more recent work (Kamide et al., in press) they found that combinatory semantics rather than simple lexical relationships influenced eye movements. For example, when viewing a carnival scene, listeners looked at a motorcycle upon hearing *The man rode…* and looked at a merry-go-round upon hearing *The girl rode…*. Thus, something higher-level than simple lexical associations influenced the pattern of eye movements.

Using a similar paradigm, Boland (2002) investigated the hypothesis that the use of a verb would implicitly introduce relevant entities (linguistic arguments) that had not yet been mentioned, and thus a picture corresponding to such an entity would draw anticipatory looks. For example, upon hearing *…mother suggested…*, participants would look at a potential recipient of the suggestion. Experiment 1 manipulated both the argument structure of the verb and the typicality/co-occurrence frequency of the target argument/adjunct, in order to distinguish between anticipatory looks to arguments specifically and anticipatory looks to pictures that were strongly associated with the verb, but did not have the linguistic status of argument. Example stimuli from Experiment 1 are

in (11). The instransitive/location stimuli provide a clear case of an adjunct target (*bed/bus*), the dative/recipient stimuli provide a clear case of an argument target (*teenager/toddler*), and the action/instrument stimuli provide an intermediate case in which the targets are arguably adjuncts (*stick/hat*). Acceptability ratings insured that sentences with typical targets were rated judged to be more acceptable than sentences with atypical targets. Furthermore, typical targets were more likely to co-occur with their verbs.[6] Importantly, there was no evidence that typical recipients had a higher co-occurrence frequency than typical locations—if anything, the opposite was true.

(11) Example stimuli from Boland (2002), Experiment 1. The typical/atypical target is underlined.

a. Intransitive/Location. *The girl slept for a while on the <u>bed/bus</u> this afternoon.* (pictures: girl, bed/bus, pillow, toy car)

b. Action/instrument. *The donkey would not move, so the farmer beat it vigorously with a <u>stick/hat</u> every day.* (pictures: donkey, farmer, stick/hat, grass)

c. Dative/Recipient. *The newspaper was difficult to read, but the mother suggested it anyway to her <u>teenager/toddler</u> last week.* (pictures: newspaper, mother, teen/toddler, dictionary)

The primary finding from Experiment 1 was that dative verbs prompted more looks to potential recipients than transitive action verbs prompted to potential instruments or intransitive verbs prompted to potential locations. This argument status effect began about 500 ms after verb onset, comparable to the timing of lexical frequency effects found by Dahan et al (2001).[7] The timing of the argument status effect suggests that it occurred during or immediately after lexical access of the verb. If verbs specify the syntactic and semantic constraints on their arguments, recognizing a verb would make available knowledge about that verb's arguments, and likely candidates (entities that

satisfy the syntactic and semantic constraints) could be identified in the current discourse model or the situational context. No overall typicality effects were found. However, toward the end of the critical interval there was some evidence of a typicality effect for instruments. By 900-1000 ms after verb onset, typical instruments were drawing more looks than atypical instruments. This late, short-lived contrast between typical and atypical instruments was the lone example of a typicality/frequency effect.

The scarcity of typicality effects in the Boland (2002) experiment must be contrasted with Kamide et al. (in press), who reported strong effects of real world knowledge in their carnival experiment. Recall that listeners looked to a motorcycle upon hearing *The man will ride…* and a to merry-go-round upon hearing *The girl will ride….* This apparent discrepancy is probably due to a difference in the way the visual stimuli were presented. In Kamide et al., argument structure knowledge introduced an abstract ride-able object into the discourse. In the visual scene, there were two ride-able objects (a motorcycle and a merry-go-round), so real world knowledge guided the viewer to the most plausible one. In Boland's experiment, the argument structure of the dative verbs introduced an abstract recipient, but there was only one potential referent pictured. In both the typical and atypical conditions, the potential referent met the lexical constraints on recipients for that particular verb, so plausibility had no opportunity to play a role. This account is consistent with prior results, indicating that pragmatic constraints influence ambiguity resolution, but not the generation of linguistic structure (Boland, 1997).

The contrast between arguments and adjuncts was replicated in Boland (2002)'s Experiment 2, in which the same animate NP (and the corresponding picture) served as an argument in the dative condition (12a) and as an adjunct in the action verb condition (12b). No instrument was mentioned in the critical trials, though a prototypical instrument for the action verb was always pictured. There were no differences in co-

occurrence frequency among the dative-recipient, active-benefactor, and active-instrument pairs.

(12) *One window was broken, so the handyman…*

[pictures: window, handyman, couple, tools]

a. ***mentioned*** *it right away to the owners****. (recipient-Argument)***

b. ***fixed*** *it hurriedly for the owners****. (benefactor-Adjunct)***

As in the first experiment, there were more looks to the target picture when it was an argument (recipient) than when it was an adjunct (benefactor, instrument) during the interval 500 - 1000 ms after the onset of the verb. There was no difference between prototypical adjunct (*fix*-tools) and improbable adjunct (*mention*-tools). Thus, the results from both experiments indicate that linguistic constraints play a privileged role. Co-occurrence frequency does not provide an alternative explanation.

The Boland data suggest an important distinction between adjuncts and arguments in terms of how verbs introduce associated entities into the discourse. More direct evidence that there are different processing mechanisms for arguments and adjuncts can be found in first pass reading data (Boland & Blodgett, submitted; Boland et al., submitted) and the SAT paradigm (Foraker & McElree, 2001).

Limitations & Open Questions in Listening Paradigms

Listening paradigms provide an exciting opportunity to investigate spoken language comprehension in an online manner. While the paradigm is largely limited to referential processing, it has been used to address a wide range of questions in word recognition, sentence analysis, and discourse processing. Because listening paradigms do not require reading, they offer the opportunity to study language processing in children and other poor readers or illiterate populations. And perhaps most importantly, listening paradigms allow, at least in principle, for investigations of language processing within conversational contexts.

Despite some clear advantages over reading paradigms, listening paradigms share the lack of an explicit model linking sentence comprehension to eye movements. In fact, this problem seems worse within listening paradigms, which haven't benefited from the same degree of scientific scrutiny over the past several decades. On the bright side, we are likely to make some progress if we examine and test our linking assumptions in a rigorous fashion. Some of the open questions include the following:

- How do directed action and passive listening tasks differ? Does the task influence the eye movements above and beyond the linguistic material?
- What cognitive events prompt a saccade to an object? How does this interact with task, local syntactic ambiguity, and other factors?
- Can we integrate models of scene perception to explain more variance in passive listening paradigms?
- How can we measure the likelihood that an object will be mentioned? How do we decide the appropriate label for an object when measuring likelihood?
- Are all of the effects essentially referential--i.e., at the level of the discourse model? Could we find a syntactic complexity effect? Could we find a lexical bias effect (e.g., ease of integrating noun/verb homophones)?

CONSIDERING READING AND LISTENING TOGETHER

Thus far, I have considered reading and listening eye movement paradigms separately, for reasons outlined in the introduction: the two approaches are very different. To the extent that they measure comparable aspects of language comprehension then, the two classes of paradigms might provide converging evidence for particular processing models or phenomena. In fact, there is growing evidence of this type of convergence in the literature.

One example of complementary reading and listening data is the argument/adjunct line of research (refs) explanation…

In other respects, convergence has not yet been achieved, but efforts toward that end are likely to lead to some important insights. For example, I have suggested that the text-based statistics commonly used to predict reading time in recent constraint-based lexicalist studies are inappropriate for evaluating anticipatory looks in listening studies. Compare the reading study of McDonald and Shillcock (this volume) to the Boland (2002) listening study described above. McDonald and Shillcock found that the higher the transitional probability of a noun, given a verb, the shorter the first fixation duration on that noun when it followed that verb in a sentence. Because they were measuring fixations on a linguistic object (a printed noun), they measured the transitional probabilities of that object, given another linguistic object (a printed verb) in text corpora. The assumption is that both the transitional probability and the fixation duration are directly related to the time it takes to access and integrate that noun.

In contrast, Boland (2002) was measuring looks to an unlabeled picture following a verb in a spoken sentence. In such a case, probability may plays some role in guiding eye movements, but it is not clear that the relevant probabilities could be obtained from a text corpus. The dependent measure assumes that the verb has been recognized, but it does not require accessing and integrating a specific word or lexical category associated with the picture because the relevant words have not yet been uttered. Rather, the dependent measure probably reflects attention to some conceptual domain that has been made salient by the pictured material and made relevant by the linguistic input thus far. Whatever the relevant measure of probability is, it should take into account the visual display as well as the linguistic material.

My point in comparing McDonald and Shillcock (this volume) with Boland (2002) is this: In reading studies, we can account for a substantial portion of the variance in reading time by assuming that the linguistic context IS the context, for all practical purposes. Furthermore, we assume that syntactic and semantic processing of some portion of text begins sometime after that text is initially fixated. However, the listening paradigm forces us to wrestle with aspects of syntactic and semantic processing that precede bottom-up word recognition, and with the fact that most real-life contexts are considerably richer than the linguistic context. While these additional layers of complexity complicate the experimental logic, they also bring some degree of real-world validity to the experiment. And it is likely that many of the insights to be gained from the listening studies will be relevant for reading studies as well.

SUMMARY

Eye movement data in both visual and auditory modalities has excelled in addressing "when" questions about how early during comprehension we have access to lexical information or how early a particular constraint is used. In this chapter, I highlighted some promising findings in each modality that push the linking assumptions between eye and mind a little further than the current state of affairs. First I suggested that, when reading unambiguous sentences, the first fixation or first pass reading time may be closely tied to syntactic generation (the access to, or construction of, possible syntactic structures). Second, I offered evidence that argument structure knowledge allows us to focus on relevant entities as soon as spoken verbs are recognized, based on anticipatory looks to objects that might be event participants. Nonetheless, neither reading nor listening paradigms may be capable of distinguishing among linguistic (and non-linguistic) levels of representation, unlike ERP's.

Acknowledgements

This chapter is based upon a presentation to the 8th Annual AMLaP conference, held in Tenerife, in September, 2002. I would like to thank the audience members and the other presenters for their comments. As I was preparing for the talk and writing this chapter, I benefited from the contributions of my collaborators, students and colleagues, especially Allison Blodgett, David Thomas, Meghan Ahern, Alicia Seifers, and the University of Michigan Psycholinguistics Research Group.

References

Ainsworth-Darnell, K., Shulman, H. G., & Boland, J. E. (1998). Dissociating brain responses to syntactic and semantic anomalies: Evidence from event-related potentials. Journal of Memory and Language, 38(1), 112-130.

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.

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

Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergrouping relational violations. Cognitive Psychology, 14, 143-177.

Boland, J. E. (1997). The relationship between syntactic and semantic processes in sentence comprehension. Language and Cognitive Processes, 12, 423-484.

Boland, J. E. (2002). Visual Arguments. Manuscript submitted for publication.

Boland, J. E., & Blodgett, A. (2001). Understanding the constraints on syntactic generation: Lexical bias and discourse congruency effects on eye movements. Journal of Memory and Language, 45(3), 391-411.

Boland, J. E., & Blodgett, A. (2002). Eye Movements as a Measure of Syntactic and Semantic Incongruity in Unambiguous Sentences. Manuscript.

Boyce, S. J., & Pollatsek, A. (1992). Identification of objects in scenes: The role of scene background in object naming. Journal of Experimental Psychology: Learning, Memory and Cognition, 18, 531-543.

Braze, D., Shankweiler, D., Ni, W., and Palumbo, L. C. (in press). Readers' eye movements distinguish anomalies of form and content. *Journal of Psycholinguistic Research*.

Brown-Schmidt, S., Campana, E., & Tanenhaus, M. K. (2002). Reference resolution in the wild: How addresses circumscribe referential domains in real time comprehension during interactive problem solving. Presented at the 15th Annual CUNY Conference on Human Sentence Processing, New York, NY.

Carlson, G. N., & Tanenhaus, M. K. (1988). Thematic roles and language comprehension. W. Wilkins (Editor), Syntax and Semantics, Volume 21: Thematic Relations . San Diego, CA: Academic Press.

Cooper, R. M. (1974). Control of eye fixation by meaning of spoken language - new

methodology for real-time investigation of speech perception, memory, and language processing. Cognitive Psychology, 6(1), 84-107.

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.

Deutsch, A., & Bentin, S. (2001). Syntactic and semantic factors in processing gender agreement in Hebrew: Evidence from ERPs and eye movements. Journal of Memory and Language, 45(2), 200-224.

Eberhard, K. M., Spivey-Knowlton, M. J., Sedivy, J. C., & Tanenhaus, M. K. (1995). Eye movements as a window into real-time spoken language comprehension in natural contexts. Journal of Psycholinguistic Research, 24, 409-436.

Ferreira, F., & Henderson, J. M. (1990). Use of verb information in syntactic parsing: Evidence from eye movements and word-by-word self-paced reading. Journal of Experimental Psychology: Learning, Memory and Cognition, 16(4), 555-568.

Fodor, J. D., & Inoue, A. (1994). The diagnosis and cure of the garden path. Journal of Psycholinguistic Research, 23, 407-434.

Fodor, J. D., & Inoue, A. (2000). Syntactic features in reanalysis: Positive and negative symptoms. Journal of Psycholinguistic Research, 29(1), 25-36.

Foraker, S., & McElree, B. (2001). The Effects of Co-Occurrence Frequency on the Time-Course of Parsing Operations. Presented at the 14th Annual CUNY Conference on Human Sentence Processing in Philadelphia, PA.

Frazier, L. (1978). On comprehending sentences: Syntactic parsing strategies. Unpublished doctoral dissertation, University of Connecticut.

Frazier, L. (1987). Theories of sentence processing. J. L. Garfield (Editor), Modularity in Knowledge Representation and Natural Language Understanding. (pp. 291-307). Cambridge, MA: MIT Press.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. Cognitive Psychology, 14(2), 178-210.

Friederici, A. D. (1995). The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. Brain and Language, 50, 259-281.

Friederici, A. D., & Frisch, S. (2000). Verb argument structure processing: The role of verb-specific and argument-specific information. Journal of Memory and Language, 43, 476-507.

Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: Effects of semantic, morphological and syntactic violations. Cognitive Brain Research, 1(3), 183-192.

Frisson, S. P., & Pickering, M. J. (1999). The processing of metonymy: Evidence from eye movements. Journal of Experimental Psychology: Learning, Memory and Cognition, 25, 1366-1383.

Garrod, S., Freudenthal, D., & Boyle, E. (1994). The role of different types of anaphor in the online resolution of sentencs in a discourse. Journal of Memory and Language, 33, 39-68.

Griffen, Z. M. & Bock, K. (2000). What the eyes say about speaking. Psychological Science, 11, 274-279.

Gunter, T., Stowe, L. A., & Mulder, G. (1997). When syntax meets semantics. Psychophysiology, 34, 660-676.

Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. Language and Cognitive Processes, 8(4), 439-483.

Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. Annual Review of Psychology, 50, 243-270.

Hollingworth, A., & Henderson, J. (1998). Does consistent scene context facilitate object perception? Journal of Experiment Psychology: General, 127, 398-415.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. Psychological Review, 87 , 329-354.

Kamide, Y., Altmann, G. T. M. & Haywood, S. (in press). Prediction and thematic information in incremental sentence processing: Evidence from anticipatory eye-movements. Journal of Memory and Language.

Mannan, S. K., Ruddock, K. H. & Wooding, D. S. (1995). Automatic control of saccadic eye movements made in visual inspection of briefly presented 2-D visual images. Spatial Vision, 9, 363-386.

Mauner, G., Tanenhaus, M. K., & Carlson, G. N. (1995). Implicit arguments in sentence processing. Journal of Memory and Language, 34, 357-382.

Mitchell, D. C. (1987). Lexical guidance in human parsing - locus and processing characteristics. Attention and Performance, 12, 601-618.

Ni, W., Crain, S., & Shankweiler, D. (1996). Sidestepping garden-paths: The contribution of syntax, semantics and plausibility in resolving ambiguities. Language and Cognitive Processes, 11, 283-334.

Ni, W., Fodor, J. D., Crain, S., & Shankweiler, D. (1998). Anomaly detection: Eye movement patterns. Journal of Psycholinguistic Research, 27(5), 515-539.

O'Brien, E. J., Shank, D. M., Myers, J. L., & Rayner, K. (1988). Elaborative Inferences: Do they occur on-line? Journal of Exeperimental Psychology: Learning, Memory, and Cognition, 3, 410-420.

Osterhout, L., & Nicol, J. (1999). On the distinctiveness, independence and time course of the brain responses to syntactic and semantic anomalies. Language and Cognitive Processes, 14(3), 283-317.

Pearlmutter, N. J., Garnsey, S. M., & Bock, K. (1999). Agreement processes in sentence comprehension. Journal of Memory and Language, 41, 427-456.

Pickering, M. J., & Traxler, M. J. (1998). Plausibility and recovery from garden paths: An eye-tracking study. Journal of Experimental Psychology: Learning, Memory, and Cognition, 24, 940-961.

Pollatsek, A., & Rayner, K. (1990). Eye movements and lexical access in reading. D. A. Balota, G. B. Flores D'Arcais, & K. Rayner (Editors), Comprehension processes in reading (pp. 143-163). Hillsdale, NJ: Lawrence Erlbaum Associates.

Rayner, K. (1983). The perceptual span and eye movement control during reading. K. Rayner (Editor), Eye movements in reading: Perceptual and language processes . New York: Academic Press.

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

Rayner, K., Reichle, E. D., & Pollatsek, A. (1998). Eye movement control in reading: An overview and model. G. Underwood (Editor), Eye guidance in reading and scene perception (pp. 243-268). New York: Elsevier.

Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, R. A., & Clifton, C. Jr. (1989). Eye movements and on-line language comprehension processes. Language and Cognitive Processes, 4, 21-50.

Rayner, K., Sereno, S. C., & Raney, G. E. (1996). Eye movement control in reading: A comparison of two types of models. Journal of Experimental Psychology: Human Perception and Performance, 22, 1188-1200.

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. Psychological Review, 105, 125-157.

Reilly, R. G., & O'Regan, J. K. (1998). Eye movement control during reading: A simulation of some word-targetting strategies. Vision Research, 38, 303-317.

Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving

incremental semantic interpretation through contextual representation. Cognition, 71(2), 109-147.

Spivey-Knowlton, M., & Sedivy, J. (1995). Resolving attachment ambiguities with multiple constraints. Cognition, 55, 227-267.

Spivey-Knowlton, M., & Tanenhaus, M. K. (1998). Syntactic ambiguity resolution in discourse: Modeling effects of referential context and lexical frequency within an integration-competition network. Journal of Experimental Psychology: Learning, Memory and Cognition, 24, 1521-1543.

Sussman, R. S., Campana, E., Tanenhaus, M. K., & Carlson, G. M. (2002). Verb-based access to instrument roles: Evidence from eye movements. Poster presented at the 8th Annual Architectures and Mechanisms of Language Processing Conference, Tenerife, Canary Islands, Spain.

Tanenhaus, M. K., & Trueswell, J. C. (1995). Sentence comprehension. J. Miller, & P. Eimas (Editor), Handbook of perception and cognition: Speech, language & communication (Vol. 11pp. 217-262). San Diego, CA: Academic Press.

Traxler, M. (2002). Plausibility and subcategorization preference in children's processing of temporarily ambiguous sentences: Evidence from self-paced reading. Quarterly Journal of Experimental Psychology Section A - Human Experimental Psychology, 55(1), 75-96.

Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999a). The development of on-line language comprehension abilities in children. Language, Minds, & Brains: Studies in Languages, 34, 209-215.

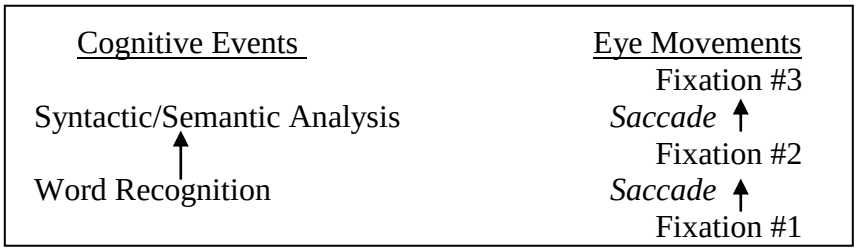
Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999b). The kindergarten-path effect: Studying on-line sentence processing in young children. Cognition, 73, 89-134.

Table 1. Anomaly Detection Results (Local First Pass Measures)

Experiment	Measure	Subcategory	Semantic	Morpho-Syntactic
Boland & Blodgett, 2002	READING TIME	X		
	% Regression	X		X
Ni et al. 1998; Braze et al., in press	READING TIME			
	% Regression			X
Pearlmutter et al. 1999, Exp 2	READING TIME			
	% Regression			X
Deutsch & Bentin, 2001	READING TIME			X
Frisson & Pickering, 1999	READING TIME		X	

Experiment	Measure	Subcategory	Semantic	Morpho-Syntactic
Boland & Blodgett, 2002	READING TIME	X		
	% Regression	X		X
Ni et al. 1998; Braze et al., in press	READING TIME			
	% Regression			X
Pearlmutter et al. 1999, Exp 2	READING TIME			
	% Regression			X
Deutsch & Bentin, 2001	READING TIME			X
Frisson & Pickering, 1999	READING TIME		X	

Figure 1. Linking Comprehension to Behavior



[1] Some readers will be bothered by the absence of a comma in these sentences. However, similar garden path effects are found in other structures without "the comma problem."

[2] Although the eye movement record can be analyzed in different ways, it is important to keep in mind that we don't have four or five **independent** dependent measures. There are really just two: fixation duration and saccade direction. Everything else is computed from that. So the fact that we see the same pattern in probability of regression, regression path time, and total time is not a finding of note. It would be quite surprising if we did not.

[3] In fact, Frederici et al. generates the prediction that both types of anomalies should be "late", given that the phrase level violations in the Boland and Blodgett stimuli rely upon argument structure knowledge. Frederici et al. assumed a rule-driven parser that did not utilize detailed lexical knowledge.

[4] For present purposes, this is not considered a local first pass measure because it includes time spent in other regions.

[5] In fact, Griffen and Bock (2000) compared a picture inspection task with a picture description task, using simple line drawings with two participants (e.g., a mouse squirting a turtle). They found an equal number of looks to the two participants during the first 1300 ms of the picture inspection. In contrast, people who were describing the pictures tended to look at the first mentioned participant during this same time interval. Thus, the nature of the experimental task can clearly influence looking patterns during the initial seconds of picture viewing.

[6] Counting co-occurrence frequency in an "anticipatory looking" paradigm presents an interesting problem. The dependent measure is an anticipatory look to a photograph, prior to hearing the target phrase. Thus, the frequency measures that are often used in parsing/reading research are inappropriate. In reading studies, the co-occurrence frequency between a verb and a phrase of a particular class (e.g., a PP beginning with *to* or a phrase that is assigned a particular thematic role) is often used to predict processing difficulty for a phrase of the same class (e.g., Boland et al., 2002; Garnsey et al., 1997; Novick & Trueswell, 2001). In contrast, what we need to know here is, given a particular verb, how likely is the occurrence of an object or person like the one in the target picture. Boland (2002) used two different methods to compute co-occurrence frequency, but it is likely that this research problem will require continued attention.

[77] When no competitor items were pictured (Experiment 2), Dahan et al. (2001) found that high frequency items like *bed* were fixated about 563 ms after onset of the word, versus 625 ms for a low frequency target like *bell* following the instruction *Point to the bed/bell*.