Chapter 15
Constraint Satisfaction Accounts of Lexical and Sentence Comprehension

Maryellen C. MacDonald and Mark S. Seidenberg

Consider what it takes to understand an ordinary sentence such as *The man bought a tie with tiny white diamonds.* Part of your understanding includes that *the man* is the agent of the action *bought* and that *a tie* is the thing being bought. To get this far, you also need to understand that *man* and *tie* are nouns and not verbs, although the verb usage of these words is possible in other contexts, such as *man the boats* or *tie your shoes*. You also need to understand that *a tie* refers to neckware, not to a game with equal scores, and that *tiny white diamonds* are an attribute of the tie, and not the currency used to purchase the tie (cf. *The man bought a tie with his credit card*). Despite these and many other possibilities where interpretation could go wrong, the odds are in favor of your interpreting this sentence correctly. For example, *man* is more common as a noun than a verb, so a comprehender who unconsciously goes with the best odds will get to the right interpretation here. Similarly, words that follow determiners such as *the* and *a* are far more likely to be nouns than verbs, and tiny white diamonds are unlikely to be offered in trade for haberdashery, at least in most cultures. Comprehenders who follow the most likely alternatives will get to the correct interpretations of these aspects of the sentence. The idea that language comprehension is a process of following likely alternatives to derive an interpretation of ambiguous input forms the basic claims of constraint satisfaction, or constraint-based, theories of language comprehension. As in these examples, what is a likely alternative depends on properties of both individual bits of information (e.g., the frequency with which a word is used as a noun or verb) and combinations of bits of information (e.g., *the* + *man* or *a* + *tie)*. Constraint-based theories emphasize how people learn, represent, and use such probabilistic information. This chapter will provide an overview of this approach, including its history, how it compares to alternative views, and a description of the kinds of computational mechanisms that are thought to underlie learning and using such constraints.

1. TRADITIONAL VIEWS OF LEXICAL AND SYNTACTIC AMBIGUITY

As the sentence about the man and his new tie illustrated, ambiguity is ubiquitous in language. This chapter will focus on two main types of ambiguity: *lexical ambiguity*, illustrated by the multiple meanings of words such as *tie*, and syntactic ambiguity,

Handbook of Psycholinguistics: 2nd Edition
ISBN: 0-12-369374-8

illustrated by the alternative interpretations of *with tiny white diamonds* as something describing the tie (thus modifying a noun) or describing the method of buying (modifying a verb). (See the Pickering and van Gompel, Kluender, and Tanenhaus and Trueswell chapters for discussions of other types of ambiguity.) The two kinds of ambiguity can interact; for example, adopting the noun vs. verb interpretation of *man* affects how one interprets the syntactic structure of a sentence containing this word. Despite the close relationship between these two types of ambiguity, for much of the history of modern psycholinguistics they have been studied independently. This division reflected differing views about lexical and syntactic representations (MacDonald, Pearlmutter, & Seidenberg, 1994). The meanings and other properties of words have often been thought to be stored in the *lexicon*, a person's mental dictionary. On this view, interpreting words involves looking up, or *accessing*, information in the lexicon. This process was thought to be autonomous, proceeding in the same way regardless of the context in which a word occurred (Tanenhaus, Leiman, & Seidenberg, 1979; Swinney, 1979). It was also thought to make minimal demands on limited capacity working memory and attentional resources, allowing multiple meanings of words to be accessed in parallel. This led to a two-stage model of lexical ambiguity resolution. In the first stage, the lexical system accessed the common meaning or meanings of words; in the second stage, information derived from the linguistic and extra-linguistic contexts and the comprehender's background knowledge were used to select the appropriate meaning and integrate it into the developing representation of the sentence (see Simpson, 1981, for review).

Syntactic structures, in contrast, were standardly assumed to be *constructed* by a mental parser on the basis of grammatical rules. Deriving sentence structure was assumed to place demands on working memory and attentional resources that are limited in capacity (Frazier, 1987; Gibson, 1998; Just & Carpenter, 1992; MacDonald, Just, & Carpenter, 1992). These memory demands caused the parser to pursue only a single interpretation of syntactic structure at a time. This also led to a two-stage model. In the first stage, general parsing principles were used to assign a candidate syntactic structure online; in the second stage, other types of knowledge were utilized to flesh out this representation (e.g., interpret it semantically) and to revise the initial analysis if it were discovered to be incorrect.

Both lexical and syntactic accounts were motivated in part by appeals to the notion that language consists of distinct modules involving different types of information and processes; however, in the lexical case, this resulted in multiple alternatives being considered in a parallel process, whereas in the syntactic case, this resulted in a single interpretation being considered in a serial process (see MacDonald et al., 1994; Tanenhaus & Trueswell, 1995, for reviews). The modular view was consistent with distinctions between the lexicon and syntax in grammatical theories that were prominent at the time the two-stage accounts were being developed (Newmeyer, 1980). The two-stage approach was also justified on the basis of assumptions about processing capacity limitations and the need to analyze the linguistic input very rapidly. The route to efficient interpretation was thought to be via a two-stage system in which the preliminary first-stage analysis prevented the input from being lost from working memory; the burden on working memory limitations was reduced because processing at this stage was limited to certain types of information, e.g., syntactic structure. The initial interpretation could then be refined, corrected, and elaborated in the second-stage analysis.[1] This attention to the time pressures of language comprehension and to the notion that processing may proceed through several distinct stages was reflected in the use of behavioral measures that were closely time-locked to the language input (e.g., tracking eye-movements, cross-modal priming). In various forms, two-stage approaches formed the dominant theoretical framework for word and sentence comprehension through the 1980s, and the focus on the time course of processing continues to this day.

The alternative view, which came to be called constraint-based language comprehension (or comprehension via probabilistic constraints), emerged in the 1990s. This approach challenged essentially every major tenet of the two-stage accounts. Whereas the two-stage theories held that comprehension consists of discrete stages at which different types of information and processes are used, constraint-based theories viewed comprehension as continuous and homogeneous, with the same types of information and processes in use at all times. Whereas the two-stage theories assumed that processing limitations restrict the types of information that initially guide the comprehension process, constraint-based theories emphasized the richness of the linguistic signal, the capacity of language users to learn this information over time, and the comprehender's capacity to bring this information rapidly to bear on the input during real-time comprehension processes.

2. SOURCES OF THE CONSTRAINT-BASED APPROACH

The constraint-based approach emerged from advances in several areas, including linguistic theory, corpus linguistics, psycholinguistics, and computational modeling.

2.1. Changing Views about Linguistic Structure

Whereas two-stage models reflect early approaches within generative grammar in which lexical and syntactic information were held to be separate, the constraint-based approach to comprehension is more closely related to work within linguistics in which (to varying degrees) lexical and syntactic representations are closely related (e.g., Bresnan, 1982; Chomsky, 1981; Joshi, 1985). The lexical representation of a word might include not only information about its spelling, pronunciation, and meaning(s) but also its grammatical functions and the types of syntactic structures in which it participates. It is a small step to then envision this information as part of a large interactive network (MacDonald et al., 1994). Under this scenario, the computation of both "lexical" and "syntactic" information in sentence comprehension is governed by a common set of lexical processing mechanisms.

---

[1] This two-stage process is reminiscent of practices in machine translation, in which an automatic but limited first pass analysis (by machine) is then corrected and elaborated in a second-stage analysis (by a human translator).

2.2. Changing Views about Context

Language is comprehended essentially as it is perceived (Marslen-Wilson, 1975), and so a central question is what types of information can be brought to bear on decoding and interpreting the incoming signal. Studies of the role of language context in comprehension have also undergone a significant shift over the years. Research in the two-stage era focused on the use of real-world knowledge in guiding the comprehension process, and on the difficulties inherent in accessing relevant information online (Kintsch & Van Dijk, 1978). People know a vast amount about the world; as research on natural language processing in artificial intelligence suggested, it is a difficult problem to design a comprehension system that accesses relevant information from an enormous database of facts (Hayes-Roth & Jacobstein, 1994). Moreover, several studies emphasized the ineffectiveness of context, suggesting that comprehenders were limited in their application of real-world knowledge during comprehension (Forster, 1979), that context facilitated lexical processing only when words were highly predictable (Fischler & Bloom, 1979), and that this very strong degree of contextual constraint is rare in naturally occurring texts (Gough, Alford, & Holly-Wilcox, 1981). These results led to the conclusion that context-based prediction was not an important component of comprehension.

Complementary findings emerged from the study of lexical ambiguity resolution (e.g., Swinney, 1979; Tanenhaus et al., 1979). Many words are ambiguous between semantically distinct meanings (e.g., WATCH: a timepiece, to look; BANK: a monetary institution, the ground bordering a river). These early studies examined the processing of ambiguous words for which there are two main meanings that are used approximately equally often in the language ("equibiased" ambiguities). The main finding was that subjects initially activated multiple meanings, even in contexts that were highly disambiguating. For example, the contexts in (1) and (2) clearly disambiguate the word ROSE. Yet subjects showed priming (facilitation compared to an unrelated control) for target words related to both of the main meanings (e.g., FLOWER, STOOD) presented immediately following each sentence (Tanenhaus et al., 1979). Results such as this were taken as evidence that comprehenders initially activated the common meanings of ambiguous words and within about 250 ms selected the correct meaning based on the context. Here too the processing of words seemed to be independent from processes involved in integrating a sequence of words into a meaningful, syntactically structured representation.

1. They all ROSE.
2. He bought a ROSE.

The research on predictability effects and lexical ambiguity resolution led many researchers to conclude that context effects are relatively weak, with the result that theories instead emphasized bottom-up aspects of processing – how words are identified. The ambiguity research played an important role in Fodor's (1983) development of his concept of modularity. The lexicon was seen as a paradigmatic example of an autonomous module in the comprehension system.

Subsequent research has led many of these conclusions to be revised. Whereas the word predictability studies initially argued for a limited role of context, later work suggested that context effects could operate at levels other than predicting specific words. Studies of semantic priming, for example, suggested that the processing of a word is facilitated when preceded by a word with which it shares semantic features (e.g., McRae, de Sa, & Seidenberg, 1997). Here the target words are not predictable, but facilitation occurs nonetheless. Moreover, it is worth noting that most of the studies which suggested that context effects are limited in scope examined reading rather than spoken language. Written language does not exhibit many of the properties that make speech perception such a difficult computational problem (e.g., variability with respect to rate, pitch, accent; relatively lower signal – noise ratios; co-articulation and the absence of definitive markers for phoneme or word boundaries). The spoken code seems inherently more context bound, insofar as the mere perception of sounds depends on the contexts in which they occur (e.g., Samuel & Pitt, 2003).

As with the context research, the lexical ambiguity research was similarly reexamined. Whereas initial studies had argued for activation of multiple meanings of ambiguous words independent of context, subsequent research yielded a more complex picture. Several studies showed that contextual information could result in only one meaning of an ambiguous word being considered online (e.g., Simpson & Kreuger, 1991). However, other studies showed that context could not override all aspects of lexical knowledge, in particular the relative frequencies of the meanings: there was still an ambiguity effect (computation of multiple meanings) when contexts favored the less-frequent meaning of an ambiguous word (Duffy, Morris, & Rayner, 1988). Thus, the system is apparently neither strictly modular nor completely context-bound. Kawamoto (1993) developed a computational model that provided insight about these results. His system was not inherently modular, insofar as nothing architectural prohibited contextual information from affecting meaning activation. However, in practice lexical information was activated more rapidly, limiting the effects of context. This is because there is a much closer relationship between the spelling or sound of a word such as ROSE and its meanings than there usually is between either of the meanings and the contexts in which they occur.

Finally, researchers began to question a key assumption underlying much of the research on lexical ambiguity: that words have discrete meanings that can be accessed like entries in a mental dictionary. The meaning of a word routinely shifts as a function of the context in which it occurs. Consider a word such as *piano*. It has a seemingly simple, unambiguous common meaning: large keyboard instrument with steel wires struck by felt-covered hammers (we are ignoring here the secondary musicological sense meaning "soft in volume"). Yet different shades of meaning are involved in pushing a piano (where weight is relevant but musical properties are not) vs. playing a piano (where the opposite is true; Merrill, Sperber, & McCauley, 1981). How to properly characterize meanings is a difficult issue that has been addressed from many theoretical and disciplinary perspectives (Margolis & Laurence, 1999). Here it is sufficient to note that it may be an essential property of word meaning that it is computed in a context-dependent manner every time a word is comprehended. This type of computation seems inherently at odds with a modular lexicon that automatically and independently activates stored meanings and passes them along to other comprehension systems. The creation of novel meanings from proper nouns (Clark & Clark, 1977) and the interpretation of novel noun compounds (Gagne & Shoben, 1997) raise similar issues.

We do not have a general theory of lexical ambiguity resolution in hand; to have one would be to solve a good part of the problem of language comprehension. However, this research made it clear that a broad range of factors involving properties of both words and contexts affect lexical ambiguity resolution, and that the interactions among these many factors determine the outcomes that are observed.

2.3. Changing Views about Language Statistics

Languages exhibit statistical structure – variations in the distributions of elements such as sounds, words, and phrases. Despite the existence of this structure, for many years statistical analyses of language attracted little interest within mainstream linguistics and psycholinguistics, principally because Chomsky (1957) compellingly argued that language exhibits important properties that are not captured by mere statistics (as "Colorless green ideas sleep furiously" illustrated). According to the probabilistic constraints approach, however, comprehension essentially *is* the process of exploiting statistical regularities of many kinds. Learning and using language seem like difficult problems (ones that necessarily require innate grammatical knowledge, or learning or parsing mechanisms) only because this statistical information was systematically excluded from theorizing.

The ground-breaking studies that expanded notions about the range of information that might be used in sentence comprehension were Bever (1970) and Ford, Bresnan, and Kaplan (1982). In a classic article, Bever (1970) made a number of important observations concerning syntactic complexity and ambiguity and the factors that can make sentence comprehension difficult. Bever suggested that comprehenders are guided by *perceptual strategies* that assign interpretations based on frequency and plausibility. He described a specific strategy whereby comprehenders interpret noun-verb-noun sequences as agent-action-object. Violating this expectation (as in Bever's example "The horse raced past the barn fell") creates a misanalysis, which came to be known as a "garden path" effect (Frazier, 1978). Ford, Bresnan, and Kaplan (1982) provided an early investigation of the effects of lexical knowledge on sentence comprehension. They proposed that comprehenders initially adopt an analysis of a syntactic ambiguity that incorporates the most frequent subcategorization of the sentence's verb (see also Fodor, 1978). Verb subcategorization refers to the noun phrase arguments a verb may take; for example *move* may or may not have a direct object noun phrase. Ford et al. provided evidence consistent with the idea that the several subcategorization options were ordered by frequency, and that comprehenders consider sentence interpretations in the corresponding order.

Although their importance was widely recognized, the Bever and Ford et al. articles did not immediately generate a program of research. One problem that inhibited further progress was that the research tools that were available did not make it easy to calculate robust language statistics from large samples of text or discourse. This problem was largely obviated in the 1990s, when resources such as the *Wall St. Journal* corpus (Marcus, Santorini, & Marcinkiewicz, 1993) became publicly available and could be analyzed using desktop computers. This methodological advance made it possible to conduct behavioral studies examining the use of various types of statistical information in comprehension (discussed below). A second problem was the absence of a theory that could explain which language statistics are relevant, and how they could be learned, represented in memory, and efficiently used in processing. In the absence of such a theory, it was not obvious how the Bever and Ford et al. findings could be extended. This problem also began to be addressed in the 1990s, with advances in the theory of statistical learning within the connectionist framework, to which we now turn.

2.4. Development of the Connectionist Paradigm

The term "connectionism" refers to a broad, varied set of ideas, loosely connected (so to speak) by an emphasis on the notion that complexity, at different grain sizes or scales ranging from neurons to overt behavior, emerges from the aggregate behavior of large networks of simple processing units. Our focus is on the parallel distributed processing (PDP) variety developed by Rumelhart, McClelland, Hinton and others in the 1980s (McClelland, Rumelhart, & Hinton, 1986). This approach includes a variety of concepts that are potentially relevant to language. In brief, PDP networks consist of large numbers of simple processing units that take on activation values. The connections between units carry weights that determine how activation is passed between units. The network is configured to perform a task (such as recognizing a word or object, or predicting the next word in a sentence). Learning involves gradually adjusting the weights on connections. The problem is to find a set of weights that yields performance that corresponds to human performance on the task (e.g., with respect to accuracy, generalization, developmental trajectory). Several algorithms can be employed for this purpose; they vary in how closely they mimic properties of learning at neural or behavioral levels (see Harm & Seidenberg, 2004, for discussion). Network performance is determined by several main factors: (1) the architecture of the system (e.g., the configuration of units and connections); (2) the characteristics of the input and output representations; (3) characteristics of the patterns used in training the model; and (4) characteristics of the learning algorithm. In other words, the model's performance depends on its initial state, what it experiences, and how it learns from those experiences.

This theoretical framework has been discussed extensively elsewhere; here we focus on three properties that inform the probabilistic constraints approach to comprehension.

First, the networks incorporate a theory of statistical learning. The main idea is that one way that people learn (there may be others) is by gathering information about the frequencies and distributions of environmental events. This type of learning is thought to be general rather than language specific. Many nonhuman species are also capable of rudimentary forms of statistical learning (Estes, 1955); humans may be distinct with respect to the power of their statistical learning capacities. Language, for example, requires tracking correlations and covariation across multiple types of linguistic information within and across modalities (e.g., a speech signal and the context in which it is uttered).

The applicability of these ideas to language was initially explored in the context of learning inflectional morphology (Rumelhart & McClelland's, 1986, past tense model) and learning to read (Seidenberg & McClelland, 1989; Plaut McClelland, Seidenberg, & Patterson, 1996; Harm & Seidenberg, 2004). The reading models in particular developed the idea that lexical knowledge consists of statistical relations between orthographic, phonological, and semantic codes. Learning then involves acquiring this statistical knowledge over time. Subsequent research on statistical learning in infants and adults has provided strong evidence consistent with this view. A wealth of studies now attest to humans' robust abilities to learn statistical patterns that inhere in diverse types of stimuli (Saffran & Sahni, in press). The domain-generality of statistical learning is suggested by studies showing that infants are equally good at learning the statistical structure in a series of spoken syllables and a series of pure tones (Saffran, Aslin, & Newport, 1996; Saffran, Johnson, Aslin, & Newport, 1999), and by similarities across auditory (Saffran et al., 1996) and visual (Kirkham, Slemmer, & Johnson, 2002) modalities. This learning mechanism provides a way to derive regularities from relatively noisy data, a property that is likely to be highly relevant to the child's experience in learning language. Although some researchers have argued that specifically grammatical relationships are not acquired by statistical learning (e.g., Marcus et al., 1999; Peña, Bonatti, Nespor, & Mehler, 2002), these claims have been challenged (Perruchet, Tyler, Galland, & Peereman, 2004; Seidenberg, MacDonald, & Saffran, 2002).

Second, the models provide a basis for understanding why particular types of statistics are relevant and not others. Above we described the main factors that determine a model's behavior (and, by hypothesis, a person's). Note that this description did not include a specification of which *types* of statistics a model should compute. It is not necessary to stipulate this in advance because this aspect of a model's behavior falls out of the other factors. In practice, what a model learns is heavily determined by the nature of the representations that are employed. These representations (e.g., of phonology or semantics) are intended as (simplified) claims about what people know and bring to a task such as language learning. This knowledge may be innate or may itself be learned by processes to be explored in other models. The goal is to endow a model with exactly the knowledge and capacities that people (infants, children, adults) bring to learning a task, although this ideal is only approximated in any implemented model. Given the properties of these representations, other aspects of the model architecture (e.g., number of units or layers; patterns of connectivity between layers), and a connectionist learning algorithm, the model will pick up on particular statistical regularities implicit in the examples on which the model is trained. Thus, motivating the various elements of a model and how it is trained is very important, but the model itself determines which statistics are computed.

This discussion is relevant to a concern that is often voiced about connectionist models, that they are too powerful – capable of learning regularities that humans cannot learn. In fact, what such models learn is highly constrained. Constraints on what is learned arise not so much from the learning algorithm itself as from other aspects of the network, particularly properties of the representations that are used. For example, models that represent articulatory or acoustic primitives in a realistic way are constrained by facts about what people can say or hear.

Third, the framework provides a powerful processing mechanism, the exploitation of multiple simultaneous probabilistic constraints. Information in a network is encoded by the weights. The weights determine ("constrain") the output that is computed in performing a task. Processing involves computing the output that satisfies these constraints. This output changes depending on what is presented as input (e.g., the current word being processed in a network that comprehends sentences word by word).

This type of processing, known as constraint satisfaction, has several interesting properties. One is that the network's output is determined by all of the weights. Such models illustrate how a large number of constraints can be utilized simultaneously without imposing excessive demands on memory or attention. Constraint satisfaction is passive – activation spreads through a network modulated by the weights on connections – rather than a resource-limited active search process. Another important property is that the constraints combine in a nonlinear manner. Bits of information that are not very informative in isolation become highly informative when taken with other bits of information. Much of the power and efficiency of the language comprehension system arises from this property. Languages exhibit many partial regularities. Different types of information are correlated, but weakly. The comprehender cannot wholly rely on any one type of information, but combinations of these partial cues are highly reliable. This concept may seem paradoxical at first. If individual cues are unreliable, wouldn't combinations of these cues be even more unreliable? No, not if cue combination is nonlinear. The informativeness of each cue varies as a function of other cues. This point is easy to grasp by illustration. Someone is thinking of an object – guess what it is. The cues are it is a fruit, it is yellow, and its name begins with B. In isolation, each cue only weakly constrains the answer. The combination of cues, however, makes it very likely that the object is *banana*.

The same process can occur on a sentence or discourse level. In the context of a discussion of shopping and the syntactic environment of the determiner *a*, the word *tie* probably refers to neckwear. This contingency holds despite the fact that all the simple probabilities are quite low – by itself, a shopping context doesn't demand that neckware be discussed, the occurrence of *a* does not predict the word *tie*, and *tie* in isolation affords several more frequent interpretations than the neckware one.

The bases of constraint satisfaction systems have been explored extensively in the computational literature. Connectionist models provide one way of implementing this process, but there are symbolic systems that perform similarly (Mackworth, 1977). In the psycholinguistic literature, the basic idea was introduced in Bates and MacWhinney's (1989) Competition Model. Bates and MacWhinney argued that language is comprehended by following "cues" that compete with one another and are weighed as a function of their effectiveness in past comprehension events. The Competition Model incorporated the important ideas that many linguistic cues are learned and language-specific; that cues could conflict and be differentially weighed; and the importance of integrating syntactic and nonsyntactic information during comprehension. The probabilistic constraint approach can be seen as coupling many ideas embodied by the Competition Model with proposals about the statistical basis of cues ("constraints"), and how multiple constraints are learned, represented, and exploited in processing. The Competition Model has been very important in research on how children acquire knowledge of language-specific cues, how languages differ with respect to the relative prominence of different cues (e.g., word order vs. inflectional morphology), and how cue competition affects the final interpretation of a sentence. The model had less to say about the integration of many simultaneous probabilistic cues, or about online processes in comprehension (see Elman, Hare, & McRae, 2004, for discussion). Also, in the connectionist models we have described, different alternative interpretations do not directly compete. The same weights are used in processing all input patterns. The performance of the model (or person) depends on the aggregate effects of exposure to many examples. There is nothing like parallel activation of multiple alternatives, just the computation of the best-fitting output. "Competition" is realized only implicitly, because alternatives have affected the weights, not by explicitly computing and comparing alternatives.

In summary, the probabilistic constraints approach emphasizes the role of statistical information concerning the occurrence and co-occurrence of different types of linguistic and nonlinguistic information in language comprehension. Learning a language involves acquiring this information from the large sample of utterances to which every learner is exposed. The theory assumes that humans are born with (or soon develop) capacities to perceive particular kinds of information (e.g., in listening), to engage in statistical learning, and to encode what is learned in networks of neurons. Familiar types of linguistic representation such as phonemes, syllables, morphemes, words, and constituents are not represented directly in memory; rather these terms are approximate descriptions for higher level statistical generalizations that emerge with experience (e.g., Seidenberg & Gonnerman, 2000). On this view, the learner builds (or "bootstraps") a language out of statistical relations among different types of information, and skilled language comprehension involves using these statistical generalizations in processing utterances. These ideas have been extensively explored in the context of syntactic ambiguity resolution, to which we now turn.

3. PROBABILISTIC CONSTRAINTS AND SYNTACTIC AMBIGUITY RESOLUTION

Syntactic ambiguities arise when a sequence of words is compatible with more than one sentence structure. Often the syntactic ambiguity coincides with a lexical ambiguity of some sort. For example, in (3), there is an ambiguity between interpreting *Carol* as the noun phrase (NP) direct object of the verb *saw* or the beginning of a sentential complement (often termed the NP/S ambiguity). This ambiguity is linked to lexical ambiguity in the verb, which can optionally take either a direct object NP or a sentential complement. The example also illustrates another common feature of syntactic ambiguities, at least in

English, that they may be triggered by the omission of an optional word or phrase. Thus, the sentential complement sentence in (1c) could be introduced with *that* (Shanta saw that Carol …), which would remove the temporary ambiguity.

3a. Temporary ambiguity: Shanta saw Carol …
3b. NP direct object interpretation: Shanta saw [Carol], but Carol didn't see her.
3c. Sentential complement interpretation: Shanta saw [Carol would be late].

A dominant concern in syntactic ambiguity resolution has been the *timecourse* over which information is brought to bear on the ambiguity. The modular two-stage account is exemplified by Frazier and colleagues' Garden Path Model (e.g., Frazier, 1987; Frazier & Clifton, 1996), in which the first-stage *parser* (the syntactic interpretation component of the comprehension system) develops a syntactic structure for the input, guided by only the lexical categories of the input words (*noun, verb,* etc.), the syntactic rules of the language, and by structure-based heuristics (Minimal Attachment and Late Closure) that direct structure building when more than one alternative structure is afforded by the input. At some later point, a second stage integrates semantic and contextual information into the representation, and if this information conflicts with the initial interpretation built by the parser, the conflict may trigger a revision and reanalysis of the input.

The constraint-based view argues that the preference for one interpretation over another during comprehension of an ambiguous sentence stems not from global heuristics such as Minimal Attachment but from the rapid combination of many probabilistic constraints. A key observation concerning such constraints is that different types of information tend to be correlated; for example, a verb's meaning is strongly related to the kinds of noun arguments it tends to appear with in sentences (Hare, McRae, & Elman, 2003; Levin, 1993; Roland & Jurafsky, 2002). As a result, even weak cues can combine with other correlated cues and have a strong effect on interpretation preferences. Thus the approach links syntactic level information, such as knowledge about transitive sentence structures (those with a direct object in the verb phrase), to lexically specific information, such as the frequency with which a particular verb (*bought,* say) occurs with a direct object, the frequency with which a noun (e.g., *tie*) occurs as a direct object, and the conjoint frequency with which *bought* and *tie* co-occur in a verb/direct object configuration. The correlation of cues has an important role in understanding how abstract pragmatic constraints, often thought to be too complex to be brought to bear in online ambiguity resolution, could have a rapid effect on the process. For example, new entities introduced into a discourse are more likely to receive additional modification than are previously mentioned (or "given") noun phrases, thus affecting the probability that syntactically ambiguous prepositional phrase will modify this noun phrase. The given/new distinction is strongly correlated with the type of determiner used to introduce the noun phrase; new entities often occur with *a*, and given ones with *the*. Thus *a tie with* is more likely to have the *with* phrase modify the tie than is the sequence *the tie with…* (Spivey-Knowlton & Sedivy, 1995). This pattern reflects how discourses are structured and might require extensive computation in some cases, but the comprehender has a ready proxy in the simple co-occurrence of some determiners and the interpretation of *with*.

The contrast between two-stage and constraint-based accounts has often focused on the extent to which the separate processing stages posited by two-stage models are isolable. In the case of the Garden Path model, in which a purely syntactic first stage is followed by use of all other types of information in a second stage, the issue is the extent to which putatively second-stage nonsyntactic information could be shown to affect the operations of the first-stage parser. A significant body of work in the 1980s and 1990s used eye fixations during reading to address this issue, and a number of researchers suggested that the earliest eye fixations on a small region of text reflected operations of the first-stage parser, while later fixations were driven by second-stage semantic integration processes (e.g., Rayner, Carlson, & Frazier, 1983). This view was motivated in part by studies in which manipulations of semantic information in syntactically ambiguous sentences were found to affect late eye fixations, but not early ones (e.g., Ferreira & Clifton, 1986; Rayner et al., 1983). Subsequent studies suggested that the delayed effects of nonsyntactic information in these reading patterns were attributable to weak or infelicitous contexts or other biases in the ambiguous stimuli (Altmann & Steedman, 1988; Crain & Steedman, 1985). More robust manipulations of context have shown clear evidence of the use of nonsyntactic information in first pass reading measures (e.g., Garnsey, Pearlmutter, Meyers, & Lotocky, 1997; Trueswell, Tanenhaus, & Garnesy, 1994) and even in the first fixation durations on words (Speer & Clifton, 1998), a measure that has often been taken as the earliest processing evidence obtainable with eyetracking (Rayner et al., 1983). Moreover, as the nature of contextual effects received additional investigation, the number of potentially relevant constraints, and the interactions between them, grew more complex. This trend can be illustrated by considering one particular ambiguity in detail.

3.1. An Example: Main Verb vs. Reduced Relative Ambiguities

The structures considered here are probably the most thoroughly studied in psycholinguistics. The focus on these structures arose from Bever's (1970) observation that whereas the sentence *The horse raced past the barn fell* is taken to be gibberish by most speakers of English, it is readily comprehended when two optional words (a relative pronoun and a form of be, such as *that was*) are inserted marking the start of a relative clause, as in *The horse that was raced past the barn fell*. Another example, somewhat easier to comprehend, is given in (4). This is called the Main Verb/Reduced Relative (MV/RR) ambiguity because it is initially unclear whether the first verb, *raced* in Bever's example and *arrested* in (4), is the main verb of the sentence (as in 4b) or is introducing a reduced relative clause (4c). The clause is said to be "reduced" because of the omission of the optional relative pronoun and a form of *be*.

4a. Temporary Main Verb/Reduced Relative Ambiguity: The three men arrested...
4b. Main Verb Interpretation: The three men arrested the burglary suspects in a parking garage.
4c. Reduced Relative Interpretation: The three men arrested by the local police were wanted in connection with the jewel robbery.

Early studies of interpretation of this ambiguity manipulated the degree of contextual information consistent with the reduced relative (RR) interpretation and found strong misinterpretation or "garden-path" effects in reading patterns at all levels of contextual support, indicated by long reading times in the sentence region that disambiguated the ambiguity (Rayner et al., 1983; Ferreira & Clifton, 1986). These reading patterns were taken to indicate that comprehenders initially adopted the main verb (MV) interpretation (the one favored by the parsing heuristic Minimal Attachment) independent of context and were surprised when the disambiguation favored the reduced relative interpretation. Subsequent studies explored the nature of contexts in depth and suggested that interpretation is guided by a number of probabilistic constraints, with the difficulty of a given interpretation of the ambiguity typically a function of several constraints acting together. Some of the major categories of constraints are listed below.

A. *Animacy of the pre-verbal NP* (e.g., *men*), as this affects the likelihood that this noun will be the agent vs. patient of an upcoming verb, in that animate nouns are more typical agents. This constraint is important because the noun is the agent of the next verb in the MV interpretation, and it is the patient of the verb in the RR interpretation (Trueswell et al., 1994; but Ferreira & Clifton, 1986 failed to find animacy effects).
B. *The relative frequency of usage of the ambiguous verb* (e.g., *arrested*) in active vs. passive structures, as the MV interpretation is an active structure while the RR is a passive. Active/passive voice frequency is related to several intercorrelated properties of the verb, including the verb's frequency of occurrence in the past tense (required for the active MV interpretation) vs. past participle (required for passives and the RR interpretation), and its relative frequency of uses in transitive (with a direct object) vs. intransitive (no direct object) constructions. The RR interpretation is always transitive, but the MV may be intransitive (Hare, Tanenhaus, & McRae, 2006; MacDonald, 1994; MacDonald et al., 1994; Trueswell, 1996).
C. *Plausibility of the pre-verbal NP as an agent vs. patient of the ambiguous verb*, such as the plausibility that *men* would be the agent vs. patient of *arrested*, what McRae, Ferretti, & Amyote (1997) called *thematic fit*. It is an example of a *combinatorial* constraint, in that it integrates properties of at least two words and the information in constraints A-B above (McRae, Spivey-Knowlton, & Tanenhaus, 1998; Pearlmutter & MacDonald, 1992; Tabossi, Spivey-Knowlton, McRae, & Tanenhaus, 1994). The power of such combinatorial constraints can be seen in several reanalyses of failures rapid effects of some simple constraint. For example, studies that found to find only minimal effects of noun animacy or other broader discourse plausibility factors (Ferreira & Clifton, 1986; Rayner et al., 1983) tended to use stimuli with verbs that strongly promoted an active, intransitive interpretation. In this situation, verb biases were working strongly in favor of the MV interpretation, and combinatorial constraints (over properties of both verbs and nouns) had little effect (MacDonald et al., 1994; Tanenhaus & Trueswell, 1995; Trueswell et al., 1994).
D. *The basic frequency of the MV vs. RR structure*. Within two-stage models, the initial preference for MV structures stems from parsing heuristics such as Minimal Attachment, but within the constraint-based tradition, this effect emerges from the

fact that MV structures are more common than RR structures in the language (Bever, 1970; McRae et al., 1998).

E. *The nature of the words after the onset of the ambiguity*. In most empirical studies, the first few words after the ambiguous verb constitute a prepositional phrase, such as *by the local police* in sentence (4c) above. Depending on their lexical properties and that of the preceding material, the post-ambiguity words may serve to promote one or the other interpretation of the ambiguity. The constraints here can be simple, such as the basic probability that *by* refers to an agent of an action (promoting the passive and thus an RR interpretation) vs. a location (less constraining for the two alternative interpretations), or the constraints may be combinatorial, such as properties of *by* given a particular preceding verb or NP, as in *by + the local police* (MacDonald, 1994; McRae et al., 1998). Following the prepositional phrase, the relative clause typically ends in most stimulus materials, and the true main verb of the sentence is encountered, as in *were wanted* in (4c). Researchers often assume that encountering the main verb completely disambiguates the string in favor of the RR interpretation, but the degree of disambiguation actually varies greatly with particular stimulus items. The major factor here is whether the main verb is itself ambiguous between a past tense and a past participle interpretation. A tense ambiguity at the main verb permits a second temporary MV/RR ambiguity in the stimulus sentence, as in *The witness examined by the lawyer turned out to be unreliable* (from Ferriera & Clifton, 1986). Here *turned* initially permits a RR modifying *lawyer* (as in *the lawyer turned in by the detective*), so that a definitive disambiguation is delayed. Stimuli with this second ambiguity are rare in most studies (including in Ferreira & Clifton), but they may be a source of additional noise in the reading data in some experiments. This additional ambiguity also serves to reinforce the point about the large number of constraints that can influence ambiguity resolution here.

F. *The thematic role of the pre-verbal noun*. Relative clauses are more natural when the head noun is a theme of the action (*the flowers* (*that were*) *sent to the performer…*) than when the recipient of the action is the head noun (*the performer (who was) sent the flowers…*) (Keenan & Comrie, 1977). The rarity (or oddness) of modifying a Goal role decreases the likelihood of a reduce relative interpretation, and some studies that have found poor use of nonsyntactic constraints have tended to contain stimuli in which the goal role is relativized (e.g., Rayner et al., 1983), which strongly promotes the MV interpretation. This bias also interacts with the effect of post-verbal words described in point E, in that when the goal NP is modified as in *the performer sent the flowers*, the words after the ambiguous verb (*the flowers*) can strongly promote the MV interpretation (Tabor, Galantucci, & Richardson, 2004).

G. *Constraints from the broader discourse that could promote either interpretation*s. These constraints include whether the discourse makes it plausible to modify the first noun, which promotes an RR interpretation (Altmann & Steedman, 1988; Crain & Steedman, 1985; Ni, Crain, & Shankweiler, 1996; Sedivy, 2002), whether the tense of the verbs in prior discourse promotes interpretation of the ambiguous verb as a past tense or past participle (Trueswell & Tanenhaus, 1991), and factors affecting the likelihood of using a reduced vs. unreduced relative clause form in various discourse situations (McKoon & Ratcliff, 2003). The influence of these discourse-level constraints may be modulated by more robust lexical-level constraints. For example, Filik, Paterson, and Liversedge (2005) found the extent to which attention-focusing words such as *only* influenced ambiguity resolution (Ni et al., 1996; Sedivy, 2002) varied with the range of alternative interpretations permitted by the ambiguous verb.

This and similar lists of potential constraints and their interactions (Townsend & Bever, 2001; Tanenhaus, Spivey-Knowlton, & Hanna, 2000) suggest why comprehension of the MV/RR ambiguity can sometimes succeed easily and other times fail miserably. Successful comprehension occurs when a variety of constraints strongly promote the ultimately correct reduced relative interpretation at an early point in the ambiguous string, and garden-pathing occurs when the evidence strongly favors the incorrect interpretation, as in *The horse raced past the barn fell*.

3.2. Computational Modeling of Constraint-Based Theory

The probabilistic constraints approach draws heavily on computational concepts such as constraint satisfaction. One way of exploring such concepts is by implementing simulation models. Modeling is a tool that can be used for different purposes; here we describe three ways computational models have been used in the development of the constraint-based approach.

1. As a tool for developing and illustrating novel theoretical concepts and analyses of the comprehension process. These models tend to be narrow in scope and tied to phenomena rather than the results of particular behavioral studies. Perhaps the most influential example is the work of Elman (1990), who developed the concept of a simple recurrent network in which the task is to predict the next word in a sentence given the current word and information about the prior context (in Elman's networks, the state of the hidden unit layer). Elman's models exhibited several interesting behaviors: they learned to predict words that were grammatical continuations of sentences; they formed representations of the grammatical categories of words; they encoded long-distance dependencies, not merely transition probabilities between adjacent words. The models introduced the important idea that sentence comprehension could be construed as following a trajectory in the state-space defined by a recurrent network. The models also helped to revive the idea that prediction might be an important component of sentence interpretation; recall that early results suggesting that words are not generally predictable from context (Gough et al., 1981) led to the view that contextual effects were weak and unhelpful. However, Elman's networks, and other connectionist approaches that emphasized distributed representations, suggested that comprehension processes might incorporate partial predictions where expectations are generated for certain semantic or syntactic properties of the upcoming input, even if exact words themselves are not predicted. There is now increasing evidence that human comprehension processes incorporate these predictive elements (e.g., Altmann, van Nice, Garnham, & Henstra, 1998; McRae, Hare, Elman, & Ferretti, 2005).

The St. John and McClelland (1990) model combined an Elman network with a distributed hidden layer of sentence meaning representations (the "sentence gestalt"). The model took simple sentences as input and was trained to answer queries about the thematic roles of noun phrases. The fact that the model was trained to represent meaning rather than predict upcoming words made it an interesting departure from other models of sentence processing, but it was very limited in scope, having a small number of words and thematic roles. The model also could not interpret multi-clause sentences, a serious limitation given the centrality of these constructions in syntactic and psycholinguistic research. Rohde (2002) adapted and expanded the St. John and McClelland architecture in a much larger model. His model replicated several key results in human sentence interpretation, but also behaved in some ways that differed from humans. Further research would be needed to determine whether these limitations could be overcome within this architecture.

A final example is Allen and Seidenberg's (1999) model in which both comprehension and production were simulated within a single network. They used the model to illustrate a theory of how people make grammaticality judgments, and how, paradoxically, this ability could be maintained in aphasia (Linebarger, Schwartz, & Saffran, 1983). The model also illustrated why a sentence such as "Colorless green ideas sleep furiously" is judged grammatical even though the transition probabilities between words are low.

None of these models are "complete" in any sense, and the limitations on their scope allow the possibility that the results are not general. Nonetheless, these models are important as vehicles for introducing novel mechanisms, approaches, and analyses.

2. As a procedure for discovering the statistical regularities implicit in a large corpus of utterances, as discussed above. This application of modeling has been used mainly in studies of language acquisition; studies by Mintz (2003), Redington, Chater, and Finch (1998) and others show how representations of grammatical categories can be derived from distributional information. Cassidy, Kelly, and Sharoni (1999) used a simple feedforward network as a procedure for discovering phonological correlates of proper names. Haskell, MacDonald, and Seidenberg (2003) used a network to discover phonological properties associated with adjectival modifiers in English. Mirković, MacDonald, and Seidenberg (2005) developed a model that learned much of the complex inflectional system for nouns in Serbian, and showed that gender was cued by correlations between phonology and semantics.

3. As a way of accounting for behavioral data. This usage is akin to models of word reading that simulate the results of behavioral experiments (e.g., Harm & Seidenberg, 2004). Some simple recurrent networks (SRNs) have been used for this purpose. These SRNs are typically subject to additional analyses that allow their behavior to be linked to measures of human performance (Christiansen & Chater, 1999; MacDonald & Christiansen, 2002; Tabor, Juliano, & Tanenhaus, 1997; Tabor & Tanenhaus, 1999). For example, MacDonald and Christiansen computed an error

measure, called Grammatical Prediction Error, that provided a good fit with reading times in studies of relative clause comprehension. Tabor, Juliano, & Tanenhaus, (1997) and Tabor and Tanenhaus (1999) related model performance to reading times using a different approach. They coupled an SRN and a dynamical processor that was designed to relate hidden unit representations to a sentence interpretation. The dynamical processor's ability to settle on an interpretation varies with past experience, and processor time can be related to reading times in behavioral studies.

Another class of models directly addresses the process of constraint-weighing during comprehension and the linkage between these processes and behavioral data (Elman et al., 2004; McRae et al., 1998; Spivey & Tanenhaus, 1998). These models are not simulations of sentence comprehension per se but rather test claims for activation of alternative interpretations at different points in a sentence. Moreover, the models have no learning component but are instead hand-tuned to the properties of particular stimulus sentences from one or more existing behavioral study, potentially yielding a transparent relationship between the activation levels in the model and the pattern of behavioral data. For example, McRae et al. modeled the interaction of six constraints in the MV/RR ambiguity. They developed stimulus materials for an empirical study and used a combination of corpus analyses and questionnaire studies to assess the degree of constraint promoting MV vs. RR interpretations for each of their stimulus items. The constraints were implemented in a simple localist model in which each constraint and each interpretation of the ambiguity are represented by single nodes in a network. The model simulates the timecourse of constraint interaction because constraints are combined at each word position, and the alternative interpretations receive activation as a function of the combined constraint strength at that point. The alternative interpretations compete with one another, so that activation of one alternative drives down activation of the other. McRae et al. used the model to test two alternative accounts of ambiguity resolution for this structure by varying the time at which different constraints were available to the model. In one model, all syntactic and nonsyntactic constraints were available as soon as the relevant words were encountered in the sentence; this model corresponded to the claim that ambiguity resolution is accomplished through the rapid integration of multiple probabilistic constraints. In the second model, a syntactic constraint favoring the MV interpretation was allowed to have an early effect, and non-syntactic constraints (verb tense frequencies and the plausibility of a noun being an agent or patient of a verb) were delayed for several words. This model was designed to simulate predictions of a two-stage model, in which the MV interpretation is adopted in the first stage and use of non-syntactic information is delayed until the second stage. McRae et al. found that the more interactive model was a better fit to the data than the one in which non-syntactic constraints were delayed. They argued that the addition of modeling provides a much stronger test of alternative accounts than empirical work alone, in that the modeling effort forces commitments to particular claims about constraint interaction and its timecourse.

This brief summary serves to show that a wealth of ideas about sentence comprehension and related aspects of language have been explored using implemented connectionist models.

4. STATE OF THE SCIENCE: CONTROVERSIES, UNRESOLVED ISSUES, AND FUTURE DIRECTIONS

The approach we have described is ambitious and yet still in the early stages of development, and so there are many gaps that researchers are attempting to address. We therefore conclude this chapter by considering a series of questions.

4.1. Statistics All the Way Down?

A number of researchers have taken issue with constraint-based approaches to language comprehension processes (e.g., Frazier, 1995, 1998; Townsend & Bever, 2001). We have discussed some of these concerns elsewhere in this chapter, and our focus in this section will be on empirical studies that are designed to provide evidence that important aspects of language comprehension have a nonstatistical basis, contradicting a basic tenet of the approach. For example, McKoon and Ratcliff (2003, 2005) have challenged the constraint-based account of ambiguity resolution in the MV/RR construction, and more broadly, suggested that difficulty with RR sentences such as *The horse raced past the barn fell*, lie not in their temporary ambiguity but in the incompatibility between the construction and the verb's meaning (specifically, that *race* is a verb with internally caused changes of state). They suggest that the RR sentences and their "unreduced" counterparts, such as *The horse that was raced past the barn fell*, have subtly different meanings and uses, such that internally caused change of state verbs can appear in unreduced but not reduced relative clauses. This approach, in which certain reduced relatives are nonsensical rather than merely ambiguous, to some degree indicts all ambiguity resolution approaches to this construction. McKoon and Ratcliff's claims have been forcefully countered by McRae, Hare, and Tanenhaus (2005) and Hare et al. (in press). They trace the difficulty in McKoon and Ratcliff's examples to the frequency of passive uses of the ambiguous verbs (see constraint B in the list above), disentangle this property from meaning components, and provide additional evidence for a constraint-based account of this ambiguity.

Perhaps the most direct empirical challenge to the constraint-based accounts comes from work by Pickering, Traxler, Van Gompel, and colleagues (e.g., Traxler, Pickering, & Clifton, 1998; Van Gompel, Pickering, & Traxler, 2001; Van Gompel, Pickering, Pearson, & Liversedge, 2005), who have argued that constraint-based accounts make incorrect predictions about reading times for certain ambiguities. Specifically, they observe that constraint-based accounts predict that comprehension times should be longer for ambiguous sentences compared to unambiguous ones, owing to the fact that ambiguous sentences engender competition between alternative interpretations. In a series of studies using several different ambiguous constructions, these authors have found that globally ambiguous sentences are read more quickly, not more slowly, than unambiguous sentences. They suggest that these data argue against a constraint-based account and instead support a two-stage model in which multiple sources of information may affect a sentence's initial interpretation. However, Green and Mitchell (2006) found that the McRae et al's., (1998) computational model generally did not enter into an extended (period of competition for globally ambiguous sentences, and thus there is no prediction for longer reading times for these items relative to disambiguated ones in these models. Green and Mitchell's simulations uncover behavior in the model that runs contrary to the assumptions that Traxler et al. (1998) and Van Gompel et al. (2001, 2005) made about model performance, and their simulation results emphasize the pitfalls of relying on intuition for how an implemented model will behave. More generally, these behavioral results and simulations will serve to push alternative accounts to be more precise, especially about predictions for sentences (like globally ambiguous ones) for which people find it difficult to compute an interpretation online. Fodor (1982) and Kurtzman and MacDonald (1993) discussed the possibility that certain global quantifier scope ambiguities may never be fully resolved by comprehenders, and perhaps global syntactic ambiguities also may not always receive a definitive resolution (see also Ferreira, Ferraro, & Bailey's, 2002, account of "good-enough" sentence interpretation and brief remarks about strategic effects in reading below).

4.2. Which Statistics?

Languages exhibit many properties that can be counted; some, such as how often verbs follow nouns, seem more relevant than others, such as the frequency distribution for words in the third position in sentences. In a fully specified theory of language comprehension, it would be clear which statistical regularities people encode and use in processing, and why. Clearly, we do not have anything like that kind of theory in hand; we have some evidence about the use of particular statistics that supports the general theoretical framework. As discussed above, in principle it should not be necessary to specify "the statistics that are relevant to language" a priori because that information should fall out of an automatic procedure: a neural network (or similar formalism) that processes language, subject to constraints imposed by the architecture, representations, and input. This procedure also approximates the experience of the child, for whom the relevant statistics are learned rather than pre-specified. Above, we summarized modeling research that represents important progress toward this ideal, and some models (in limited domains) have generated testable predictions. However, there are practical limits on building large-scale models, and analyzing the behavior of a complex dynamical system becomes difficult. These conditions make it difficult to use a computational model as an independent, hands-off way of determining which statistics are relevant. For a skeptic, the absence of a complete model creates the possibility that the statistical approach is vacuous because it can explain any result. No matter how an experiment turns out, the argument goes, a researcher can find a statistic or combination of statistics that can account for the pattern of results. The approach is therefore not merely unfalsifiable (i.e., able to fit all patterns that do occur); it can also fit patterns of data that never occur.

While it is important to acknowledge the limits of current knowledge, these concerns are not realistic. First, the methodology used in this research does not involve collecting behavioral data and then finding statistics to fit the results. Rather, researchers test hypotheses developed from several sources: linguistic theory; existing empirical findings; close analyses of examples (e.g., data mining a corpus); and other types of theorizing (e.g., about why languages exhibit particular kinds of statistical regularities; see below). Second, the strategy of tailoring one's statistical analysis to fit a particular set of data would be self-defeating, because it results in overfitting: the results will not generalize to other data of the same sort (see Seidenberg & Plaut, 2006, for a discussion of this issue with respect to models of word reading). Third, the grain of the behavioral data is such that the number of constraints that actually account for detectable variance is relatively small (although not easily accommodated by factorial designs). The theory states that performance is determined by the aggregate effects of all of one's experiences with language; out of this highly variable set of data, some very strong regularities arise, and these end up being the ones that can account for observable behavior.[2] Finally, the idea that there will always be a statistic or combination of statistics to fit any data pattern greatly misjudges the extent to which language structure is constrained. There aren't unlimited degrees of freedom in accounting for the data because there aren't unlimited degrees of freedom in how a language can be structured. It is true that the number of language statistics that can be calculated is nearly infinite, but most of them are meaningless. The fact that we can calculate language statistics that do not account for data is not a problem if there are other bases for determining which statistics *are* relevant.

4.3. Different Models for Different Phenomena?

Above we suggested that implemented models have been important in developing and testing the probabilistic constraints approach. Such models are as yet limited in scope, and many important linguistic phenomena have yet to be addressed. A deeper concern is that every model is different, i.e., different models have been applied to different phenomena. Where is the integrative model that would subsume the broad range of phenomena that have as yet been investigated using many different models?

We are sympathetic to the concerns that are raised by using different models for different phenomena. It would a problem if the principles that explain one phenomenon, studied using one architecture, are incompatible with the principles that explain some other phenomenon, studied using another architecture. Perhaps such models can only succeed when narrowly focused, as might be seen if a more general model were attempted. There probably is no simple way to address this issue, or a simple, preferable alternative. In the early stages of developing this approach it has been necessary to merely demonstrate that it is sufficient to account for interesting phenomena, given the general climate of skepticism about statistical methods and connectionist models in the study of language. Further progress would be achieved if, as additional models are developed, researchers were able to identify which general computational properties

[2] Perhaps a good analogy is the analysis of evoked potentials (Kutas, van Petten, & Kluender, this volume). This methodology involves gathering many samples of apparently noisy data: every brain wave is different from every other one. Many different aspects of these waves could be measured and counted, and there is no independent theory of how the waves are generated to indicate which elements are important prior to looking. Averaging across many data samples, however, certain regularities emerge (i.e., systematic displacements of the waveform such as P300, N400 and others). Language may exhibit a greater number of regularities, and we also want a better theory of their sources, but the similarities are noteworthy.

are crucial. These principles are more important than the characteristics that differentiate implemented models. This approach has achieved some success in the domain of single word reading. Researchers have identified a small set of critical computational principles, which have been explored in a succession of models. Each model is slightly different than the others (because of advances in understanding network properties or because they focus on different phenomena), but they are governed by the same principles. It is the set of principles and how they apply to a set of phenomena that constitute the explanatory theory, not the properties of individual models.

Achieving this deeper level of understanding language comprehension requires much more research: more models addressing a broader range of phenomena; comparing different architectures with respect to the same phenomena; analyzing models to identify the properties that are critical to achieving human-like performance. This is an ambitious agenda and it is not clear whether there are enough researchers with the sufficient technical skills and interest in the approach to achieve these goals. Moreover, it is not clear whether it is either feasible or desirable to develop a genuinely integrative model of broad scope. As Seidenberg and Plaut (2006) observed,

> The concept of a complete, integrative model is a non sequitur, given the nature of the modeling methodology, particularly the need to limit the scope of a model in order (a) to gain interpretable insights from it and (b) to complete a modeling project before the modeler loses interest or dies. The goal of the enterprise, as in the rest of science, is the development of a general theory that *abstracts away from* details of the phenomena to reveal general, fundamental principles (Putnam, 1973). Each model serves to explore a part of this theory in progress.

We think it's important to keep in mind that models are tools, not the goal of the theoretical enterprise. The limitations of individual models are tolerable if they yield insights about puzzling phenomena, generate testable hypotheses, and promote theoretical development.

4.4. Where Do Language Statistics Come From?

Within the constraint satisfaction account, a fine grained characterization of the statistical regularities constraining the interpretation of ambiguities is important to capturing behavioral data. As much of the above discussion suggests, the complexity of the system makes this accounting a nontrivial enterprise. Some insight into the constraints, and a broader account of language performance, may emerge from addressing the question of the origin of the statistical regularities in language. That is, why do languages exhibit certain statistical properties and not others? At least three forces may modulate the statistics of a language. First, some statistical regularities may be shaped by conceptual structures (McRae, Ferretti, & Amyote, 1997), so that aspects of our (nonlinguistic or pre-linguistic) thinking constrains the form of utterances. Second, statistics may be shaped by language producers' sensitivity to limits on our comprehension abilities, so that producers tailor their utterances to those that are more easily understood, in the process creating statistical regularities in the language. The extent to which speakers are sensitive to listener needs is not fully resolved, but certainly there are at least some clear examples of speakers tailoring speech to their audience, such as the broad differences in the character of child- and adult-directed speech. Finally, some statistics may emerge from the production process itself. MacDonald (1999) and Gennari and MacDonald (2004) argued for this approach, termed the Production–Distribution–Comprehension (PDC) account, which suggests that certain statistical patterns emerge from language producers' needs to maximize fluency during production. For example, speakers appear to adopt syntactic structures for their utterances at least in part to yield an utterance in which more highly "accessible" words are uttered early, where accessibility here refers to a variety of conceptual, lexical, and perhaps articulatory properties that affect the ease of articulating a particular word or phrase (Bock, 1987). MacDonald (1999), Gennari and MacDonald (2004) and Race and MacDonald (2003) have applied this logic to several different comprehension issues and have argued that comrprehenders' preferences to interpret ambiguities in favor of one vs. another alternative structure can be linked to the relative frequency of those alternatives in the language, owing to speakers' and writers' syntactic choices during the production process. These choices in turn stem from biases inherent in the production system, such as to place shorter sentence elements (words or phrases) before longer ones, or to place pauses or small optional function words before sections of high-production complexity. If this view is on the right track, then an increased understanding of constraint satisfaction in sentence comprehension will emerge from a better grasp of how the production process promotes certain production choices (word orders, word-structure co-occurrences, structure-discourse co-occurrences, etc.) and discourages others.

4.5. Where To Go Next?

In presenting this approach, we have already mentioned several important directions for future research. We will close by mentioning three more. First, as detailed in the chapter by Trueswell and Tanenhaus in this volume, researchers are beginning to expand the range of constraints that comprehenders consider by investigating the extent to which comprehenders integrate the visual scene and other aspects of conversational interaction. This work allows an investigation of comprehension of speech, in contrast to the vast majority of studies discussed in this chapter, which have investigated written language. Second, returning to the written language realm, something that would benefit all theoretical perspectives is to increase our understanding of reading data and its relationship to computational accounts of comprehension processes. Researchers from many theoretical perspectives agree that the theorizing and the data are not well matched, in that certain reading patterns are compatible with radically different interpretations of ambiguity resolution processes (e.g., Lewis, 2000; Tanenhaus, 2004; Van Gompel et al., 2001). This situation may be traced to some combination of imprecision in theoretical claims, inability of reading or other dependent measures to resolve fine-grained predictions about timecourse, insufficient consideration of the possibility that reading and other dependent measures may reflect comprehenders' strategies so that the data may not be a "pure" reflection of the ambiguity resolution processes. That is, we all know that a novel, a newspaper, and a chemistry textbook elicit different reading behaviors, yet there is very little appreciation among researchers of the degree to which reading strategies might vary with the nature of filler items or comprehension questions in experiments and the extent to which these strategic components could be affecting reading patterns that are attributed to "automatic" sentence processing operations.

Third, the constraint-based approach affords the opportunity to investigate the relationship between acquisition and skilled performance. The focus in adult comprehension has been on timecourse, specifically the speed with which comprehenders can bring constraints to bear on linguistic input, and there has been relatively little discussion of the learning mechanisms by which comprehenders come to possess the relevant constraints. The claim that the learning is inherently statistical invites research into the extent to which there is continuity between acquisition and adult performance and the extent to which a statistical learning account will prove adequate to explain the child's rapid mastery of language. These questions link to an enormous and ongoing research enterprise in child language acquisition, one with its own controversies and struggles to match theory and data. It is therefore an exciting possibility for theoretical development that the studies of the adult state and the acquisition process in the child might be mutually informative and constraining.

REFERENCES

Allen, J., & Seidenberg, M. S. (1999). The emergence of grammaticality in connectionist nets. In: B. MacWhinney (Ed.), *The Emergence of Language.* Hillsdale, NJ: Erlbaum.

Altmann, G. T. M., van Nice, K., Garnham, A., & Henstra, J. A. (1998). Late closure in context. *Journal of Memory and Language*, *38*, 459–484.

Altmann, G. T. M., & Steedman, M. J. (1988). Interaction with context during human sentence processing. *Cognition*, *30*, 191–238.

Bates, E., & MacWhinney, B. (1989). Functionalism and the competition model. In: B. MacWhinney, & E. Bates (Eds), *The crosslinguistic study of sentence processing.* Cambridge, UK: Cambridge University Press.

Bever, T. G. (1970). The cognitive basis for linguistic structure. In: J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 279–360). New York: Wiley.

Bock, J. K. (1987). An effect of the accessibility of word forms on sentence structure. *Journal of Memory and Language*, *26*, 119–137.

Bresnan, J. (1982). The passive in lexical theory. In: J. Bresnan (Ed.), *The mental representation of grammatical relations* (pp. 3–86). Cambridge, MA: MIT Press.

Cassidy, K. W., Kelly, M. H., & Sharoni, L. J. (1999). Inferring gender from name phonology. *J. Exp. Psychology: General*, *128*, 362–381.

Chomsky, N. (1957). *Syntactic Structures.* The Hague: Mouton.

Chomsky, N. (1981). Knowledge of language: Its elements and origins. *Philosophical Transactions of the Royal Society of London*, *295*, 223–234.

Christiansen, M. H., & Chater, N. (1999). Toward a connectionist model of recursion in human linguistic performance. *Cognitive Science*, *23*, 157–205.

Clark, H. H., & Clark, E. V. (1977). *Psychology and language*. *An introduction to psycholinguistics*, New York: Harcourt Brace Jovanovich.

Crain, S., & Steedman, M. (1985). On not being led up the garden path: The use of context by the psychological parser. In: D. Dowty, L. Karttunen, & A. Zwicky (Eds), *Natural Language Parsing*. Cambridge: Cambridge University Press.

Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language*, *27*, 429–446.

Elman, J. L. (1990). Representation and structure in connectionist models. In: G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 345–382). Cambridge, MA: MIT Press.

Elman, J., Hare, M., & McRae, K. (2004). Cues, constraints, and competition in sentence processing. In: M. Tomasello, & D. Slobin (Eds), *Beyond nature-nurture: Essays in honor of Elizabeth Bates.* Mahwah, NJ: Lawrence Erlbaum Associates.

Estes, W. K. (1955). Statistical theory of distributional phenomena in learning. *Psychological Review*, *62*, 369–377.

Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language*, *25*, 348–368.

Ferreira, F., Ferraro, V., & Bailey, K. G. D. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, *11*, 11–15.

Filik, R., Paterson, K. B., & Liversedge, S. P. (2005). Parsing with focus particles in context: Eye movements during the processing of relative clause ambiguities. *Journal of Memory and Language*, *53*, 473–495.

Fischler, I. S., & Bloom, P. A. (1979). Automatic and attentional processes in the effects of sentence contexts on word recognition. *Journal of Verbal Learning and Verbal Behavior*, *18*, 1–20.

Fodor, J. A. (1983). *The Modularity of Mind.* Cambridge, MA: MIT Press.

Fodor, J. D. (1978). Parsing strategies and constraints on transformations. *Linguistic Inquiry*, *9*, 427–473.

Fodor, J. D. (1982). The mental representation of quantifiers. In: S. Peters, & E. Saarinen (Eds), *Processes, beliefs, and questions* (pp. 129–164). Dordrecht: Reidel.

Ford, M., Bresnan, J., & Kaplan R. M. (1982). A competence-based theory of syntactic closure. In: J. Bresnan (Ed.), *The mental representation of grammatical relations*, Cambridge, Massachusetts, MIT Press, pp. 727–796.

Forster, K. I. (1979). Levels of processing and the structure of the language processor. In: W. Cooper, & E. C. T. Walker (Eds), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett* (pp. 27–85). Mahwah, NJ: Lawrence Erlbaum Associates.

Frazier, L. (1978). *On comprehending sentences: Syntactic parsing strategies*. Ph.D. thesis, University of Connecticut.

Frazier, L. (1987). Sentence processing: A tutorial review. In: M. Coltheart (Ed.), *Attention and performance 12: The psychology of reading* (pp. 559–586). Mahwah, NJ: Lawrence Erlbaum Associates.

Frazier, L. (1995). Constraint satisfaction as a theory of sentence processing. Journal of Psycholinguistic Research, 24, 437–468.

Frazier, L. (1998). Getting there (slowly). Journal of Psycholinguistic Research, 27, 123–146.

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

Gagne, C. L., & Shoben, E. J. (1997). Influence of thematic relations on the comprehension of modifier-noun combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(1), 71–87.

Garnsey, S. M., Pearlmutter, N. J., Myers, E., & Lotocky, M. A. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language*, 37, 58–93.

Gennari, S., & MacDonald, M. C. (2004). Relating comprehension and production of relative clauses. Paper presented at the 17th Annual CUNY conference on human sentence processing, College Park, MD.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, *68*, 1–76.

Gough, P. B., Alford, J. A., & Holley-Wilcox, P. (1981). Words and contexts. In: O. J. L. Tzeng, & H. Singer (Eds), *Perception of print: Reading research in experimental psychology* (pp. 85–102). Mahwah, NJ: Lawrence Erlbaum Associates.

Green, M. J., & Mitchell, D. C. (2006). Absence of real evidence against competition during syntactic ambiguity resolution, *Journal of Memory and Language*, 1–17.

Hare, M., McRae, K., & Elman, J. L. (2003). Sense and structure: Meaning as a determinant of verb subcategorization references. *Journal of Memory and Language*, *48*, 281–303.

Hare, M., Tanenhaus, M. K., & McRae, K. (2006). Understanding and producing the reduced relative construction: Evidence from ratings, editing and corpora. Manuscript submitted for publication.

Harm, M., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review*, *111*, 662–720.

Haskell, T., MacDonald, M. C., & Seidenberg, M. S. (2003). Language learning and innateness: Some implications of *compounds research*. *Cognitive Psychology*, *47*, 119–163.

Hayes-Roth, F., & Jacobstein, N. (1994). The state of knowledge-based systems. *Communications of the ACM*, *37*, 26–39.

Joshi, A. K. (1985). How much context-sensitivity is required to provide reasonable structural descriptions: Tree adjoining grammars. In: D. Dowty, L. Kartunnen, & A. Zwicky (Eds), *Natural Language Parsing*. Cambridge, NY: Cambridge University Press.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*, 122–149.

Kawamoto, A. H. (1993). Nonlinear dynamics in the resolution of lexical ambiguity: A parallel distributed processing account. *Journal of Memory and Language*, *32*, 474–516.

Keenan, E. L., & Comrie, B. (1977). Noun phrase accessibility and universal grammar: *Linguistic Inquiry*, *8*, 63–99.

Kelly, M. (1992). Using sound to solve syntactic problems: The role of phonology in grammatical category assignment. *Psychological Review*, *99*, 349–364.

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, *85*, 363–394.

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence of a domain general learning mechanism. *Cognition*, *83*, B35–B42.

Kurtzman, H., & MacDonald, M. C. (1993). Resolution of quantifier scope ambiguities. *Cognition*, *48*, 243–279.

Kutas, M., Van Petten, C., & Kluender, R. (this volume). Psycholinguistics electrified II: 1995–2005. In: M. Traxler, & M. A. Gernsbacher (Eds), Handbook of psycholinguistics (2nd ed.). New York: Elsevier.

Levin, B. (1993). *English verb classes and alternations: A preliminary investigation.* Chicago, IL: The University of Chicago Press.

Lewis, R. L. (2000). Falsifying serial and parallel parsing models: Empirical conundrums and an overlooked paradigm. *Journal of Psycholinguistic Research*, *29*, 241–248.

Linebarger, M., Schwartz, M., & Saffran, E. (1983). Sensitivity to grammatical structure in so-called agrammatic aphasics. *Cognition*, *13*, 361–393.

MacDonald, M. C. (1994). Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes*, *9*, 157–201.

MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review*, *109*, 35–54.

MacDonald, M. C., Just, M. A., & Carpenter, P. A. (1992). Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology*, *24*, 56–98.

MacDonald, M. C. (1999). Distributional information in language comprehension, production, and acquisition: Three puzzles and a moral. In: B. MacWhinney (Ed.), *The emergence of language.* Mahwah, NJ: Erlbaum.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, *89*, 483–506.

Mackworth, A. K. (1977). Consistency in networks of relations. *Artificial Intelligence*, *8*, 99–118.

Marcus, G. F., Vijayan, S., Bandi Rao, S., & Vishton, P. M (1999). Rule learning by seven-month-old infants. *Science*, *283*, 77–80.

Marcus, M. P., Santorini, B., & Marcinkiewicz, M. (1993). Building a large annotated corpus of english: The Penn Treebank. *Computational linguistics*, *19*, 313–330.

Margolis E., & Laurence. S. (Eds). (1999). *Concepts: Core readings.* Cambridge, MA: MIT Press.

Marslen-Wilson, W. D. (1975). Sentence perception as an interactive parallel process. *Science*, *189*, 226–228.

McClelland, J. L., Rumelhart, D. E., & Hinton, G. E. (1986). The appeal of parallel distributed processing. In: D. E. Rumelhart, & J. L. McClelland (Eds), *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. I). Cambridge, MA: MIT Press.

McKoon, G., & Ratcliff, R. (2003). Meaning through syntax: Language comprehension and the reduced relative clause construction. *Psychological Review*, *110*, 490–525.

McKoon, G., & Ratcliff, R. (2005). "Meaning through syntax" in sentence production and comprehension: Reply to McRae et al. (2005). *Psychological Review*, *112*, 1032–1039.

McRae, K., de Sa, V. R., & Seidenberg, M. S. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General*, *126*, 99–130.

McRae, K., Hare, M., Elman, J. L., & Ferretti, T. R. (2005). A basis for generating expectancies for verbs from nouns. *Memory & Cognition*, *33*, 1174–1184.

McRae, K., Ferretti, T. R., & Amyote, L. (1997). Thematic roles as verb-specific concepts. *Language and Cognitive Processes*, *12*, 137–176.

McRae, K., Hare, M., & Tanenhaus, M. K. (2005). Meaning Through Syntax is insufficient to explain comprehension of sentences with reduced relative clauses: Comment on McKoon and Ratcliff (2003). *Psychological Review*, *112*, 1022–1031.

McRae, K., Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language*, *38*, 283–312.

Merrill, E. C., Sperber, R. D., & McCauley, C. (1981). Differences in semantic encoding as a function of reading comprehension skill. *Memory & Cognition*, *9*, 618–624.

Mintz, T. H. (2003). Frequent frames as a cue for grammatical categories in child directed speech. *Cognition*, *90*, 91–117.

Mirković, J., MacDonald, M. C., & Seidenberg, M. S. (2005). Where does gender come from? Evidence from a complex inflectional system. *Language and Cognitive Processes*, *20*, 139–168.

Newmeyer, F. J. (1980). *Linguistic theory in America: The first quarter-century of transformational generative grammar*. New York: Academic Press.

Ni, W., Crain, S., & Shankweiler, D. (1996). Sidestepping garden paths: Assessing the contributions of syntax, semantics and plausibility in resolving ambiguities. *Language and Cognitive Processes*, *11*, 283–334.

Pearlmutter, N. J., & MacDonald, M. C. (1992). Plausibility and syntactic ambiguity resolution. *Proceedings of the 14th Annual Conference of the Cognitive Society* (pp. 498–503). Mahwah, NJ: Lawrence Erlbaum Associates.

Peña, M., Bonatti, L. L., Nespor, M., & Mehler, J. (2002). Signal-driven computations in speech processing. *Science*, *298*, 604–607.

Perruchet, P., Tyler, M. D., Galland, N., & Peereman, R. (2004). Learning nonadjacent dependencies: No need for algebraic-like computations. *Journal of Experimental Psychology: General*, *133*, 573–583.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, *103*, 56–115.

Putnam, H. (1973). Reductionism and the nature of psychology. *Cognition*, *2*, 131–146.

Race, D. S., & MacDonald, M. C. (2003). The use of "that" in the production and comprehension of object relative clauses. *Proceedings of the 25th Annual Meeting of the Cognitive Science Society.* Boston, MA: Erlbaum.

Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning & Verbal Behavior*, *22*, 358–374.

Redington, M., Chater, N., & Finch, S. (1998). Distributional information: A powerful cue for acquiring syntactic categories. *Cognitive Science*, *22*, 425–469.

Rohde, D. L. T. (2002). A Connectionist Model of Sentence Comprehension and Production. Unpublished Ph.D. thesis, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA.

Roland, D., & Jurafsky, D. (2002). Verb sense and verb subcategorization probabilities. In: S. Stevenson, & P. Merlo (Eds), *The lexical basis of sentence processing: Formal, computational, and experimental issues* (pp. 325–346). Amsterdam: John Benjamins.

Rumelhart, D. E., & McClelland, J. L. (1986). On learning the past tenses of English verbs. In: J. L. McClelland, D. E. Rumelhart, and the PDP research group (Eds). Parallel distributed processing: Explorations in the microstructure of cognition. Volume II. Cambridge, MA: MIT Press.

Saffran, J. R., & Sahni, S. D. (in press). Learning the sounds of language. In: M. Joanisse, M. Spivey, & K. McRae (Eds), *Cambridge handbook of psycholinguistics*, Cambridge, UK: Cambridge University Press.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928.

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*, 27–52.

Samuel, A. G., & Pitt, M. A. (2003). Lexical activation (and other factors) can mediate compensation for coarticulation. *Journal of Memory and Language*, *48*, 416–434.

Sedivy, J. C. (2002). Invoking discourse-based contrast sets and resolving syntactic ambiguities. *Memory and Language*, *46*(2), 341–370.

Seidenberg, M. S., & Gonnerman, L. (2000). Explaining derivational morphology as the convergence of codes. *Trends in Cognitive Sciences*, *4*, 353–361.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*, 523–568.

Seidenberg, M. S., MacDonald, M. C., & Saffran, J. R. (2002). Does grammar start where statistics stop? *Science*, *298*, 553–554.

Seidenberg, M. S., & Plaut, D. C. (2006). Progress in understanding word reading: Data fitting versus theory building. In: S. Andrews (Ed.), *From inkmarks to ideas: Current issues in lexical processing.* Hove, UK: Psychology Press.

Simpson, G. B. (1981). Meaning dominance and semantic context in the processing of lexical ambiguity. *Journal of Verbal Learning & Verbal Behavior*, *20*, 120–136.

Simpson, G. B., & Kreuger, M. A. (1991). Selective access of homograph meanings in sentence context. *Journal of Memory and Language*, *30*, 627–643.

Speer, S. R., & Clifton, C. (1998). Plausibility and argument structure in sentence comprehension. *Memory & Cognition*, *26*(5), 965–978.

Spivey, M. J., & Tanenhaus, M. K. (1998). Syntactic ambiguity resolution in discourse: Modeling the effects of referential context and lexical frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1521–1543.

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

St. John, M. F., & McClelland, J. L. (1990). Learning and applying contextual constraints in sentence comprehension. Artificial Intelligence, 46, 217–257.

Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, *18*, 645–659.

Tabor, W., & Tanenhaus, M. K. (1999). Dynamical models of sentence processing. *Cognitive Science*, *23*, 491–515.

Tabor, W., Galantucci, B., & Richardson, D. (2004). Effects of merely local syntactic coherence on sentence processing. *Journal of Memory & Language*, *50*, 355–370.

Tabor, W., Juliano, C., & Tanenhaus, M. K. (1997). Parsing in a dynamical system: An attractor-based account of the interaction of lexical and structural constraints in sentence processing. *Language and Cognitive Processes*, *12*, 211–271.

Tabossi, P., Spivey-Knowlton, M. J., McRae, K., & Tanenhaus, M. K. (1994). Semantic effects on syntactic ambiguity resolution: Evidence for a constraint-based resolution process. In: C. Umiltà, & M. Moscovitch (Eds), *Attention and performance 15: Conscious and nonconscious information processing* (pp. 589–615). Cambridge, MA: MIT Press.

Tanenhaus, M. K. (2004). Sentence processing: Past, present, and future. In: M. Carreiras, & C. Clifton, Jr. (Eds). *On-line measures in sentence processing: Eye movements, ERPS and beyond*. Sussex, UK: Psychology Press.

Tanenhaus, M. K., Leiman, J. M., & Seidenberg, M. S. (1979). Evidence for multiple stages in the processing of ambiguous words in syntactic contexts. *Journal of Verbal Learning and Verbal Behavior*, *18*, 427–440.

Tanenhaus, M. K., Spivey-Knowlton, M. J., & Hanna, J. E. (2000). Modeling thematic and discourse context effects with a multiple constraints approach: Implications for the architecture of the language comprehension system. In: M. W. Crocker, M. Pickering, & C. J. Clifton (Eds), *Architectures and mechanisms for language processing* (pp. 90–118). Cambridge, UK: Cambridge University Press.

Tanenhaus, M. K., & Trueswell, J. C. (1995). Sentence comprehension. In: J. Miller, & P. Eimas (Eds), *Handbook of Perception and Cognition, Vol. 11: Speech, Language and Communication* (pp. 217–262). New York: Academic Press.

Townsend, D. J., & Bever, T. G. (2001). *Sentence comprehension: The integration of habits and rules.* Cambridge, MA: MIT Press.

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

Trueswell, J. C. (1996). The role of lexical frequency in syntactic ambiguity resolution. *Journal of Memory & Language*, *35*, 566–585.

Trueswell, J. C., & Tanenhaus, M. K. (1991). Tense, temporal context and syntactic ambiguity resolution. *Language and Cognitive Processes*, *6*, 303–338.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory & Language*, *33*, 285–318.

Van Gompel, R. P. G., Pickering, M. J., Pearson, J., & Liversedge, S. P. (2005). Evidence against competition during syntactic ambiguity resolution. *Journal of Memory and Language*, *52*, 284–307.

Van Gompel, R. P. G., Pickering, M. J., & Traxler, M. J. (2001). Reanalysis in sentence processing: Evidence against current constraint-based and two-stage models. *Journal of Memory and Language*, *45*, 225–258.

This page intentionally left blank

Chapter 16
Eye-Movement Control in Reading

Keith Rayner and Alexander Pollatsek

1. INTRODUCTION

Psychologists interested in language processing have increasingly turned to the use of eye movement data to examine moment-to-moment processing (Liversedge & Findlay, 2000; Rayner, 1998; Rayner & Liversedge, 2004; Starr & Rayner, 2001).[1] This, in our view, is not surprising because eye movements represent one of the best ways to study language comprehension processes. In comparison to other available techniques (see Haberlandt, 1994; Rayner & Sereno, 1994a), eye-movement data provide a relatively natural, on-line method for investigating critical psycholinguistic issues. Importantly, eye movements during reading are not part of an artificially induced task – they are part of the normal reading process. In addition, monitoring readers' eye movements does not perturb their normal reading rate. Although it has been the case in the past that studies utilizing eye-movement data have typically required the position of the head to be fixed (by the use of a bitebar or chin rest), this is not necessarily the case at the present time (although the most accurate eye-tracking data invariably result from having a participant's head fixed). This constraint is sometimes viewed as introducing an unnatural component to reading. However, our view (see also Rayner & Sereno, 1994a) is that participants in eye-movement experiments read quite normally. This is supported by data reported by Tinker (1939) that indicate that reading rate and comprehension do not differ when readers read text in a laboratory situation with their eye movements recorded and when they read in normal conditions (i.e., without a fixed head).

We do not think that the only way to study reading is by examining the eye-movement record. In building a theory of language comprehension, it is necessary to obtain converging evidence from various sources. Thus, other tasks, such as word-by-word self-paced reading, rapid serial visual presentation (RSVP) of sentences, event-related

---

[1]In this context, it is interesting to note that Table 2 in the Rayner (1998) review article lists 113 articles that were reported between 1978 and 1998 in which eye-movement data were used to examine some aspect of language processing. As of mid-2005, 88 articles have appeared in 7.5 years since the 1998 review article was published.

Handbook of Psycholinguistics: 2nd Edition
ISBN: 0-12-369374-8

potentials (ERP), and so on, have their place and many of them undoubtedly probe the nature of readers' mental representations and provide useful information about online processing of language. However, we think that among the current methodologies, the eye-movement technique does the best job of revealing moment-to-moment processes in reading.

A second general point that we wish to make is that psycholinguists interested in language processing often use eye-movement data without understanding some of the basic issues underlying the technique (see Rayner, 1998; Rayner & Liversedge, 2004). That is, it is now relatively easy to obtain eye-movement data (as many companies that market eye-trackers also provide software for data analysis), and many researchers seem to be primarily interested in seeing if their manipulation has an effect on the eye-movement record. However, we would argue that it is quite important for such researchers to know something about eye movements per se since properties of the oculomotor system could well be influencing the results obtained. In this chapter, we will provide an overview of what is known about eye movements in reading, and the relationship between eye movements and cognitive/linguistic processing. We will focus primarily on eye movements and lexical processing, though we will also touch on some research dealing with parsing and discourse processing. In large part, we will argue that the movements of the eyes through the text is primarily driven by lexical processing, with higher-order processing intervening when something does not compute well. We will also describe some recent models of eye-movement control in reading, though we will focus primarily on our own model (the E–Z Reader model).

2. BASIC FACTS ABOUT EYE MOVEMENTS IN READING

During reading, we typically have the impression that our eyes are gliding smoothly across the page. However, this is an incorrect impression; instead the eyes make a series of rapid movements (called *saccades*) separated by periods of time when the eyes are relatively still (called *fixations*). It is only during the fixations that new visual information is encoded from the text because vision is functionally suppressed during the saccades. Fixations typically last about 200–250 ms, although individual fixations in reading can be as short as 50–100 ms and as long as 500 ms. Distributions of fixation durations look like normal distributions (with the mean around 200–250 ms) that are skewed to the right. Typically, saccades last roughly 20–40 ms; the duration of the saccade depends almost exclusively on the size of the saccade. Saccades moving from the end of one line to the next (called *return sweeps*) typically last longer than the movements that progress along a line, and they also tend to undershoot the intended target. Thus, a return sweep will often be followed by a corrective movement to the left (when reading English). Nevertheless, the first fixation on the line is typically 5–7 letter spaces from the beginning letter on the line; likewise, the last fixation on a line is also typically 5–7 letters from the last letter in the line. Thus, only about 80% of the text typically falls between the extreme fixations.

While the two eyes begin moving at about the same time, it turns out that the eyes do not land in exactly the same place in a word. Liversedge, White, Findlay, and Rayner

(2006; see also Heller & Radach, 1999) recently demonstrated that in as much as 50% of the cases, the two eyes are not aligned on the same character. Nevertheless, despite this divergence (which does vary as a function of line location), the effect of linguistic processing is still apparent in the eye-movement record (Juhasz, Liversedge, White, & Rayner, 2006).

The saccades per se serve the function of bringing a given region of text into foveal vision for detailed analysis. A given line of text falling on the reader's retina can be divided into three different regions with respect to the reader's point of fixation: foveal, parafoveal, and peripheral. The foveal region corresponds to the central 2° of visual angle around the fixation point (for text at a normal viewing distance, 1° of visual angle is equivalent to roughly 3–4 letters); the fovea is specialized for processing detail. The parafoveal region of a line extends from the foveal region out to about 5° of visual angle to each side of fixation. Readers are able to acquire some useful letter information from this region (see Section 3.1). The peripheral region includes everything on the line beyond the parafoveal region. Beyond the fovea, acuity drops off markedly and words that are not located in the fovea are difficult to identify. Indeed, reading on the basis of non-foveal information is difficult if there is parafoveal information and impossible if only peripheral information is available (Rayner & Bertera, 1979). Although readers are aware of the ends of lines and other gross aspects of the text, information in peripheral vision tends to be of little use in reading.

The average saccade size in reading is about 7–9 letter spaces.[2] However, just as with fixation durations, there is quite a bit of variability in saccade size: some saccades are as short as one letter and some can be over 20 letter spaces (though the longest saccades typically follow a regression and take the eyes to a point ahead of the point at which the regression was launched). The variability that exists in both fixation duration and saccade size is related to processing activities: when text is difficult, readers make longer fixations and shorter saccades. Furthermore, when text is difficult, readers move their eyes backwards in the text (these backwards movements are called *regressions*). Regressions occur about 10% of the time in skilled readers. Many regressions are short (back to the word just to the left of the current fixation) and probably reflect either oculomotor irregularities or else word recognition difficulties; other regressions are longer, and probably reflect comprehension difficulties. Interestingly, there also appears to be an inhibition of return component to regressions as fixations preceding saccades to previously fixated words are longer than saccades to skipped words (Rayner, Juhasz, Ashby, & Clifton, 2003).

Much of the research on eye movements and reading has focused on fixation time on a word (or on reading time for larger segments of text). However, both the probability of a regression from a word and the probability of skipping a word are often examined as well.

[2]In reading, the appropriate metric for how far readers move their eyes is character spaces and not visual angle (see Morrison & Rayner, 1981).

Part of the variability in saccade length discussed above is related to word skipping (generally, skipped words are processed even though they are not fixated). Skipping is not random, as short words (three or fewer letters) are skipped fairly frequently, six-letter words are usually fixated, and words that are eight letters or longer are rarely skipped (Brysbaert & Vitu, 1998; Rayner & McConkie, 1976). Other factors influence skipping as content words are typically fixated about 83% of the time, whereas function words (which obviously tend to be shorter) are fixated only 19–38% of the time (Carpenter & Just, 1983; Rayner & Duffy, 1988). (We will return to factors that influence skipping below.)

While eye-movement data are very informative with respect to lexical processing and understanding reading, they are not perfect reflections of the mental activities associated with comprehension. There is a purely motoric component of eye movements, and low-level visual and oculomotor factors can also influence fixation time and saccade length. Nevertheless, very useful information can still be obtained from the eye-movement record.

3. CRITICAL ISSUES IN USING EYE-MOVEMENT DATA TO STUDY READING

If one is interested in using eye-movement data to study some aspect of language comprehension during reading, there are a number of issues inherent in using eye movements that need to be addressed. We will briefly discuss the following issues: the perceptual span, integration of information across saccades, control of eye movements, and measures of processing time.

3.1. The Perceptual Span

How much information do readers process on each fixation? What is the size of the effective field of view? These questions are clearly related to issues of acuity that we discussed above. Clearly, for readers of English, most of the useful information is confined to the foveal and parafoveal regions. Indeed, studies by McConkie and Rayner (1975), Rayner (1975), and Rayner and Bertera (1979) using gaze-contingent display change paradigms have confirmed this. With these techniques, either the global amount of information available to the reader can be precisely controlled (as in the moving window paradigm, McConkie & Rayner, 1975; Rayner & Bertera, 1979), or the amount and type of information in a specific region can be precisely controlled, as when a preview word is changed to a target word mid-saccade (as in the boundary paradigm, Rayner, 1975). These experiments, and many that followed (see Rayner, 1978, 1998 for reviews) have demonstrated that for readers of English (and other alphabetic writing systems), the span of perception (or region of effective vision) extends from 3–4 character spaces to the left of fixation (or the beginning of the currently fixated word) to 14–15 character spaces to the right of fixation. Futhermore, readers do not acquire useful information from lines below the one they are fixating (Pollatsek, Raney, LaGasse, & Rayner, 1993). Given that information from the rightmost part of the perceptual span is typically rather gross information, the region of word identification on the current fixation is even more restricted (to typically no more than 7–8 letters to the right of fixation, although the exact size of the region varies as a function of the text being read).

The fact that the word identification span is restricted turns out to be very advantageous for researchers interested in using eye movements to study on-line language processing. If readers could process words from a wide range around their point of fixation, it would be difficult to know which word was being processed at any point in time and eye movements would not be particularly useful for studying language processing. It would be ideal for studying language processing if readers only processed the word they were fixating (making it easy to tie down what is being processed at any point in time). The reality is not quite that good, but much of the processing when a word is fixated is on the fixated word, especially the processing that occurs before the decision to move on to the next word in the text. We will return to this issue more fully in Section 6 when we discuss models of eye-movement control in more detail.

3.2. Integration of Information across Saccades

Readers do not obtain a chunk of information on one fixation and then a different chunk of information on the next fixation. Rather, there is overlap of information from fixation to fixation. That is, they usually obtain useful information from the word to the right of the currently fixated word (and occasionally from the word two to the right) and this information is used on the following fixation. So, if a reader is looking at word n, they identify the meaning of that word, but also obtain some *preview* information from word $n+1$ that helps them identify it when they fixate it. In general, the size of this preview benefit is 30–40 ms (Hyönä, Bertram, & Pollatsek, 2004).

Research is still on-going to determine what levels of processing are responsible for this benefit from a preview of a word; however, much has been learned and we will paint the general results in rather broad strokes (see Rayner, 1998 for more precise details). First, it is clear that visual information is not integrated across fixations; if the case of the letters changes from fixation to fixation, readers do not notice the change and it has little effect on their reading (McConkie & Zola, 1979; Rayner, McConkie, & Zola, 1980). Second, semantic information is not the basis of the preview effect. Thus, *song* as a preview in the parafovea for *tune* does not result in preview benefit in reading (Altarriba, Kambe, Pollatsek, & Rayner, 2001; Rayner, Balota, & Pollatsek, 1986). In the Rayner et al. (1986) study, the same words that were ineffective as parafoveal primes produced a robust priming effect when the prime (*song*) and target (*tune*) were both presented foveally. Third, morphological information is also not a good candidate for facilitating preview benefit (Inhoff, 1989; Kambe, 2004; Lima, 1987). Fourth, letter information is important: information about the beginning letters of word $n+1$ is critically important (Rayner et al., 1980; Rayner, Well, Pollatsek, & Bertera, 1982; Inhoff, Pollatsek, Posner, & Rayner, 1987). In this latter context, it is interesting that Miller, Juhasz, and Rayner (2006) recently reported that words with early orthographic uniqueness points do not yield stronger parafoveal preview benefits than words with late orthographic uniqueness points (see also Lima & Inhoff, 1985, for further evidence that is inconsistent with COHORT types of effects in preview benefit). Further, it should be noted that information about ending letters is also important, but not as important as beginning letters (Briihl & Inhoff, 1995). More recently, Johnson, Perea, and Rayner (2006; see also Johnson, 2006) have demonstrated the importance of letter information by showing that transposed letters (*jugde* for *judge*) provide more preview benefit than letter substitutions (*jvbge* for *judge*), and almost as much benefit as identical previews. Fifth, phonological codes are important in integrating information across fixations (Ashby & Rayner, 2004; Chace, Rayner, & Well, 2005; Henderson, Dixon, Petersen, Twilley, & Ferreira, 1995; Miellet & Sparrow, 2004; Pollatsek, Lesch, Morris, & Rayner, 1992; Sparrow & Miellet, 2002). For example, in Pollatsek et al.'s study, a homophone of the target word (*beech*) presented as a preview in the parafovea (for *beach*) facilitated processing of the target word more than the preview of an orthographically related preview (*bench*). In summary, the basis for the parafoveal preview benefit effect appears to be some type of combination of abstract letter codes and phonological codes.

3.3. Eye-Movement Control

There are two components of eye-movement control: (1) what determines *where* to look next and (2) what determines *when* to move the eyes. We will discuss each in turn. But, we first want to make the point that both decisions are computed on-line on most fixations. The first unambiguous demonstration of this was provided by Rayner and Pollatsek (1981). In those experiments, the physical aspects of the text were varied randomly from fixation to fixation, and the behavior of the eyes mirrored what was seen on the current fixation. In the first experiment, the size of the window of normal text was randomly varied from fixation to fixation (so the size of the window might be 9 letters on fixation 1, 31 letters on fixation 2, 17 letters on fixation 3, and so on), and saccade length varied accordingly. In the second experiment, the foveal text was delayed after the onset of the fixation by a mask (with the time of the delay varying from fixation to fixation), and fixation durations varied accordingly (see also Morrison, 1984).

3.3.1. Where to move the eyes

Low-level information (i.e., the spaces between words) is the primary determinant of where to look next (Morris, Rayner, & Pollatsek, 1990; Pollatsek & Rayner, 1982; Rayner, Fischer, & Pollatsek, 1998). When spacing information is absent, saccades are much shorter and readers are much more cautious in moving their eyes (Rayner et al., 1998). The length of the upcoming words is also important (O'Regan, 1979, 1980; Rayner, 1979). Although there is some variability in where the eyes land on a word, readers tend to make their first fixation on a word about halfway between the beginning and the middle of the word (McConkie, Kerr, Reddix, & Zola, 1988; Rayner, 1979; Rayner, Sereno, & Raney, 1996). A recent interesting, and seemingly counterintuitive, finding is that fixations tend to be longer when readers, fixations initially land near the middle of the word than when they land on the ends of words (Vitu, McConkie, Kerr, & O'Regan, 2001). Nuthmann, Engbert, and Kliegl (2005) have elegantly demonstrated that this *inverted optimal viewing position* effect is largely attributable to mislocalized fixations.

As indicated above, word skipping is heavily influenced by word length as shorter words are more likely to be skipped (Brysbaert & Vitu, 1998; Rayner, 1998). However, linguistic variables (particularly contextual constraint, but also word frequency to some extent) also have strong influences on where decisions, notably on whether a word is skipped (though there is little influence from contextual constraint on where in the word the eyes land, Rayner, Binder, Ashby, & Pollatsek, 2001). In particular, words that are highly predictable are much more likely to be skipped (Gautier, O'Regan, & LaGargasson, 2000; Rayner & Well, 1996). More frequent words also tend to be skipped more, although this effect is not as strong as that of contextual constraint (Rayner et al., 1996). (Note that both these effects hold even when the length of the word is controlled.) Our view is that words are largely skipped because they have been identified on the prior fixation and there is some evidence suggesting that fixations prior to skips are often inflated (Drieghe, Rayner, & Pollatsek, 2005; Kliegl & Engbert, 2005; Pollatsek, Rayner, & Balota, 1986; Pynte, Kennedy, & Ducrot, 2004). Thus, whereas low-level variables are largely determining where to fixate next, if the word to the right of fixation is identified on the current fixation, such identification will lead to a change in decision about which word to target next.

3.3.2. When to move the eyes

In Section 4, we will discuss a large number of variables (related to how easy or difficult a word is to process) that have been shown to influence fixation time on a word. In this section, we will limit ourselves to discussion of how quickly information gets into the processing system and its implications for when to move the eyes.

Rayner, Inhoff, Morrison, Slowiaczek, and Bertera (1981) first demonstrated that reading proceeds essentially normally if text is presented for at least 50 ms on each fixation before a masking pattern replaces the entire text (see Ishida & Ikeda, 1989). More recently, studies by Rayner, Liversedge, White, and Vergilino-Perez (2003), Liversedge et al. (2004), and Rayner, Liversedge, and White (2005) have demonstrated that if the text is available for 60 ms prior to either the fixated word disappearing or being masked, reading proceeds quite smoothly and normally. Of greater interest is that they also found that the frequency of the fixated word has just as strong an influence on how long the eyes remain in place when it disappears after 60 ms as when it does not disappear. This appears to be a strong evidence that the cognitive processes associated with understanding the fixated word is the primary force driving the eyes through the text.

3.4. Measures of Processing Time

We will first focus on the measures most commonly used to investigate the processing time associated with a given target word. These measures are: *first-fixation duration* (the duration of the first fixation on a word), *single-fixation duration* (the duration on a word when only one fixation is made on the word), and *gaze duration* (the sum of the durations of all fixations on a word prior to moving to another word). In addition, the *total time* on a word (the sum of the durations of all fixations on a word including regressions) is often reported.

In Section 4, we will primarily discuss relevant studies in terms of the three first pass variables (first fixation, single fixation, and gaze duration), though studies dealing with specific target words typically also report the *spillover* time (typically measured as the fixation time on the word following the target word), the probability of fixating on the word, the probability of refixating the word (i.e., the probability of making additional fixations on the word following the initial fixation), and the probabilities of regressing back to the word and regressing back from the word. Another variable that has become increasingly used in studies using specific target words is *go-past time*, which is the time from first fixating on the word (including regressions back in the text) until a fixation is made to the right of it. This measure thus includes more than first pass time and can reasonably be construed as the time it takes upon reading the target word on first pass until it is successfully integrated with the on-going context.

While single-fixation duration, first-fixation duration, and gaze duration are the measures of choice for studying the time course of word recognition, a wider variety of measures is typically used in measuring processing associated with larger regions of text (as is typical in the types of studies we will discuss in Section 5). For the most part, in such studies, critical regions of text are identified, usually consisting of about 3–4 words, and the time it takes readers to read the regions of interest is measured. The standard measures are: *first-pass reading time* (the counterpart of gaze duration: the sum of all fixations in a region from first entering the region until leaving the region), *go-past* or *regression path duration* (the sum of all fixations in a region from first entering the region, including any regressions that are made, until moving to the right of the region), regressions-out (the probability of regressing out a region, generally limited to the first-pass reading of that region), second-pass reading time (the sum of all fixations in a region following the initial first pass time), and total reading time (the sum of all fixations in a region, both forward and regressive movements). First-fixation durations are also sometimes reported, especially when the region is short or when the researcher is interested in spillover effects from the previous region, but when regions are long and the disambiguating material is not likely to be included in the initial fixation, the first fixation is inappropriate.

Measures such as first pass time are generally referred to as early measures; second pass time (and total time, to the extent that it reflects second pass time rather than first pass time) are generally referred to as late measures (Rayner, Sereno, Morris, Schmauder, & Clifton, 1989). The go-past and regressions-out measures are sometimes considered early measures (but sometimes as late measures); the occurrence of a regression probably reflects some difficulty in integrating a word when it is fixated, arguably an early effect, but the operation of overcoming this difficulty may well occur late in processing. Acutally, as Clifton, Staub, and Rayner (2006) pointed out, the terms early and late may be misleading, if they are taken to line up directly with first-stage vs. second-stage processes that are assumed in some models of sentence comprehension (Rayner, Carlson, & Frazier, 1983; Frazier, 1987). Nonetheless, careful examination of when effects appear may be able to shed some light on the underlying processes. Effects that appear only in the late measures are in fact unlikely to directly reflect first-stage processes; effects that appear in the early measures may reflect processes that occur in the initial stages of sentence processing, at least if the measures have enough temporal resolving power to discriminate between distinct, fast-acting, processes.

Finally, it is ideally the case that a region of interest would consist of the same words. However, in psycholinguistic experiments this is not always possible and researchers often end up being forced to compare conditions that vary in specific words and/or the number of words. In such cases, a deviation from regression measure introduced by Ferreira and Clifton (1986) is typically used to attempt to correct (albeit imperfectly) for length differences.

4. WORD RECOGNITION AND EYE MOVEMENTS

One of the most robust findings in studies of eye movements and reading is that the ease or difficulty associated with understanding a word during reading clearly affects how long readers fixate on that word. In the remainder of this section, we will briefly review findings which have demonstrated effects due to word difficulty, contextual constraint, number of meanings (lexical ambiguity), phonological codes, semantic relations between words, morphological effects, and plausibility effects prior to moving to higher level effects. We will not provide an exhaustive review of all such studies. Rather, we will simply highlight the typical findings associated with each of these variables that plausibly have some relationship to how easy a word is to process.

4.1. Word Difficulty

There is a huge body of research on what makes individual words more or less difficult to process (in and out of context). Perhaps the most widely used standard index of word difficulty is word frequency (usually determined from corpus counts of adult reading materials). In reading, word frequency has a very reliable influence on how long readers look at a word (Just & Carpenter, 1980; Rayner, 1977). One problem in assessing the effect of word frequency is that it is fairly highly correlated with other variables, notably the length of a word. However, Rayner and Duffy (1986) and Inhoff and Rayner (1986) manipulated word frequency while controlling for word length and demonstrated that there was still a strong effect of frequency on fixation times on a word. The size of the frequency effect typically ranges from about 20 to 40 ms in first-fixation duration and from 30 to 90 ms in gaze duration (depending on the size of the difference in the actual frequencies in the stimuli). Since these initial reports, numerous studies have demonstrated frequency effects on fixation time measures (Altarriba, Kroll, Sholl, & Rayner, 1996; Calvo & Meseguer, 2002; Henderson & Ferreira, 1990, 1993; Hyönä & Olson, 1995; Kennison & Clifton, 1995; Kliegl, Grabner, Rolfs, & Engbert, 2004; Liversedge et al., 2004; Raney & Rayner, 1995; Rayner & Fischer, 1996; Rayner, Ashby, Pollatsek, & Reichle, 2004; Rayner, Fischer, & Pollatsek, 1998; Rayner, Sereno, & Raney, 1996; Sereno & Rayner, 2000; Vitu, 1991). An interesting finding is that the frequency effect is attenuated as words are repeated in a short passage (Rayner, Raney, & Pollatsek, 1995) so that by the third encounter of a high-or low-frequency word, there is no difference between the two. (The durations of fixations on both high and low frequency words decrease with repetition, but the decrease is more dramatic for low frequency words.)

Is word frequency the only variable that affects how difficult a word is to process? Obviously, one can manipulate the visibility of the letters and get sizable increases in fixation time when the letters are harder to encode (Rayner, Reichle, Stroud, Williams, & Pollatsek, 2006; Reingold & Rayner, 2006). But are there other, deeper, variables? One line of experimentation suggests that there is more to difficulty than frequency. This line of research has one set of participant's rate words as to their perceived familiarity, and then has another set read text in which target words are matched for (objective) frequency but different on rated familiarity. These experiments (Chafin, Morris, & Seely, 2001; Juhasz & Rayner, 2003; Williams & Morris, 2004) clearly demonstrated that familiarity influenced fixation times on words even when frequency was controlled (particularly for words that are low frequent). Thus, something is operative besides frequency, but then one wants to know what the objective variables are that are causing these differences in familiarity.

Another variable that has been examined in depth recently is *age-of-acquisition* (when a person is likely to have first encountered a word). Age-of-acquisition is determined either by corpus counts or by subjective ratings, and it has been shown to influence how long it takes to process a word (Juhasz, 2005). Juhasz and Rayner (2003, 2006) recently demonstrated that there was an effect of age-of-acquisition above and beyond that of frequency on fixation times in reading that was somewhat stronger than that of word frequency.

This effect raises several questions. First, is age-of-acquisition merely a cumulative frequency effect? That is, perhaps age of acquisition measures are merely better indicators of how frequently one has seen a word in text in one's lifetime than standard frequency measures. Instead, perhaps words that are learned earlier in life enjoy a special status. There is currently no resolution of this issue (see Juhasz, 2005). A second issue is whether effects such as familiarity effects are merely due to age-of-acquisition. Again, there is no clear resolution of this issue. A third issue is whether there are other variables that are confounded here. One obvious variable is the concreteness of a word, as words acquired early in life tend to be concrete and words acquired later in life tend to be abstract. Lastly, age-of-acquisition is also likely confounded with the frequency of a word in the spoken language, and given that phonological coding is important in reading (see below), this is another potentially important variable (Juhasz, 2005). The next 5 years of research will perhaps resolve these issues.

4.2. Contextual Constraint

In studies that manipulate predictability, sentence contexts are first prepared such that certain target words are either predictable or unpredictable from the context. A rating study is then performed, and the rating scores are used as the measure of predictability – how much the prior context constrains a given target word. Considerable research has demonstrated that words that are predictable from the preceding context are looked at for less time than words that are not predictable. Ehrlich and Rayner (1981) first demonstrated the effects of contextual constraint on fixation time, and the basic result has been confirmed a number of times (Ashby, Rayner, & Clifton, 2005; Balota, Pollatsek, & Rayner, 1985; Drieghe et al., 2005; Inhoff, 1984; Rayner Ashby et al., 2004; Rayner & Well, 1996; Schustack, Ehrlich, & Rayner, 1987). Not only are fixation time measures shorter on highly predictable words than low predictable words, readers also skip over highly predictable words more frequently than low predictable words (Binder, Pollatsek, & Rayner, 1999; Ehrlich & Rayner, 1981; Rayner & Well, 1996; Vitu, 1991).

One question about predictability is whether it is merely due to the objective *transitional probability* with which a given word follows another word in printed text (as determined via corpus counts). McDonald and Shillcock (2003a, 2003b) found that words with high transitional probability (e.g., *defeat* following *accept*) receive shorter fixations than words with low transitional probability (e.g., *losses* following *accept*). However, Frisson, Rayner, and Pickering (2005) subsequently found that differences in predictability were not merely due to transitional probability. In an experiment that had a highly controlled set of items, there was an effect of predictability (with transitional probability controlled), but no effect of transitional probability (with predictability controlled). This suggests that correlations between words in text have little influence unless people are conscious of them. In addition, Frisson et al. showed that predictability effects are detectable very early in the eye-movement record and between contexts that are only weakly constraining.

4.3. Number of Meanings (Lexical Ambiguity)

The number of meanings a word has influences fixation time on the word. Rayner and Duffy (1986), Duffy, Morris, and Rayner (1988), and Rayner and Frazier (1989) first demonstrated this lexical ambiguity effect, which has subsequently been replicated a number of times (Binder, 2003; Binder & Morris, 1995; Binder & Rayner, 1998; Dopkins, Morris, & Rayner, 1992; Folk & Morris, 2003; Kambe, Rayner, & Duffy, 2001; Rayner, Pacht, & Duffy, 1994; Sereno, 1995; Sereno, O'Donnell, & Rayner, 2006; Sereno, Pacht, & Rayner, 1992; Wiley & Rayner, 2000). The basic finding is that when a *balanced* ambiguous word (a word like *straw* with two approximately equally likely meanings) is encountered in a neutral context, readers look longer at it than an unambiguous control word matched on length and frequency, whereas they do not look any longer at a *biased* ambiguous word (a word like *bank* with one highly dominant meaning) in a neutral context than an unambiguous control word. In the former case, it appears that there is some sort of conflict between the two meanings. However, it appears that the subordinate meaning is not registered in the latter case; this is consistent with the finding that if a subsequent disambiguating region makes clear that the subordinate meaning was intended, then there is considerable disruption to reading (long fixations and regressions). In contrast, when the disambiguating information precedes the ambiguous word, readers do not look any longer at the balanced ambiguous word than the control word. Apparently, the context provided sufficient information to guide the reader to the contextually appropriate meaning. However, in the case of biased ambiguous words when the subordinate meaning is instantiated by the context, readers look longer at the ambiguous word than the control word. This latter effect has been termed the *subordinate bias effect*. Rayner, Cook, Juhasz, and Frazier (2006) recently demonstrated that an adjective immediately preceding the target noun is a sufficient context to produce the effect.

All of the experiments mentioned above dealt with ambiguous nouns. In this context, results reported by Frazier and Rayner (1987) and by Pickering and Frisson (2001) are quite interesting. Frazier and Rayner (1987) found that syntactic category ambiguity (*trains* can be a noun or a verb) resulted in delayed effects; fixation time differences did not emerge on the target word itself (even with biasing context), but were delayed as if the system were trying to get additional information before committing to one meaning or the other. Pickering and Frisson (2001) likewise reported that with verbs with two meanings, the resolution of verb meaning ambiguity is delayed. Frazier and Rayner (1990) also found that nouns with different senses (e.g., *newspaper* meaning a publication or a physical object) yielded delayed effects in comparison to the typically reported results with nouns with two distinct meanings, and Frisson and Pickering (1999) found that metonymic expressions were treated differently from literal expressions (as reflected by the fixation time patterns on such expressions). The reasons for the different patterns of results are, as yet, far from clear, but they all show that lexical ambiguity influences the time that it takes to process a word in text.

4.4. Phonological Coding

Words that are phonologically ambiguous (like *tear* and *wind*) have substantially longer gaze durations than unambiguous control words (Carpenter & Daneman, 1981) and words with two different spellings, but the same pronunciation (and two different meanings, such as *beech–beach* and *shoot–chute*), also have longer fixation times than unambiguous control words (Folk, 1999; Folk & Morris, 1995; Jared, Levy, & Rayner, 1999; Rayner, Pollatsek, & Binder, 1998). In addition, readers will often misinterpret the low frequency member of the pair as the higher frequency member if the context is highly constraining (Rayner et al., 1998, cf., Daneman & Reingold, 1993; Daneman, Reingold, & Davidson, 1995). Moreover, the finding previously mentioned (Pollatsek et al., 1992) that a parafoveal preview of a homophone of a target word provides greater preview benefit than a matched orthographic control indicates that phonological coding occurs early – even before a word is fixated.

More recently, Ashby and Clifton (2005) found that lexical stress influences how long readers look at a word. They further argued that implicit prosody (prosody generated internally by the reader), is a factor in how long readers look at a word (see also Hirotani, Frazier, & Rayner, 2006).

4.5. Semantic Relationships between Words

Words that are semantically related to each other (and in close proximity to each other in the text) produce effects that appear to be analogous to semantic priming (Meyer & Schvaneveldt, 1971). Thus, the word *king* in the close proximity to *queen* results in shorter fixation times on *queen* than on an unrelated word in the same location (Carroll & Slowiaczek, 1986; Morris, 1994). Carroll and Slowiaczek (1986) found that such priming effects only occurred when the two words were in the same syntactic constituent, but Morris (1994) found some evidence for priming across constituents. Morris and Folk (1998) reported that this facilitation depends in part on whether the semantic associate of the target word is in linguistic focus (see Birch & Rayner, 1997). As we noted earlier, there are also repetition effects on fixation times in reading (Rayner et al., 1995; Raney & Rayner, 1995).

Several studies have demonstrated that specific kinds of semantic processing influence reading time on a word. Traxler, McElree, Williams, and Pickering (2005) and Traxler, Pickering, and McElree (2002) investigated the effect on readers' eye movements when the context forces a noun with no intrinsic temporal component to be interpreted as an event, as in the phrase *finish the book*. They found increased go-past time on the critical word or increased first pass time on the next region (see also Frisson & Pickering, 1999). Frisson and Frazier (2005) found that when a mass noun appears with plural morphology (e.g., *some beers*) or a count noun appears in the singular with a plural determiner (e.g., *some banana*), there is an increase in the duration of the first fixation on the critical word.

4.6. Morphological Effects

Most research on word recognition has traditionally dealt with mono-morphemic words. This tradition has also been largely true of research on eye movements and word recognition. Recently, however, many studies have examined the processing of morphemically complex words (Inhoff, Radach, & Heller, 2000; Juhasz, Inhoff, & Rayner, 2005). This newer tradition started with studies (Hyönä & Pollatsek, 1998; Pollatsek, Hyönä, & Bertram, 2000) dealing with the processing of Finnish words (which by their very nature tend to be long and morphologically complex). Hyönä and Pollatsek (1998) found that the frequency of both the first and second constituent of two constituent compound words had large effects on the gaze duration on the compound word for long compound words (when the frequencies of the compound words were matched). However, Bertram and Hyönä (2003) found that the effects of the frequency of the first constituent were quite attenuated for shorter Finnish compound words. Similarly, somewhat smaller constituent frequency effects have recently been demonstrated with English compound words that were about the same length as the shorter Finnish compounds (Andrews, Miller, & Rayner, 2004; Juhasz, Starr, Inhoff, & Placke, 2003). Niswander-Klement and Pollatsek (2006) found a similar length-modulated constituent frequency effect for English prefixed words. That is, they found effects of the frequency of the root morpheme of the prefixed word (with the frequency of the words controlled), but that this effect was stronger for longer prefixed words.

Pollatsek and Hyönä (2005) recently demonstrated that semantic transparency (defined as whether the meanings of the constituents were related to the meaning of the word) had no effect on fixation times on Finnish compound words. There are conflicting findings in English, however. Juhasz (2006) found a main effect of transparency on gaze durations, whereas Frisson, Niswander-Klement, and Pollatsek (2006) obtained no effect. However, what is consistent across studies is that there is evidence for morphological decomposition for both opaque and transparent compounds, as there is an effect of the frequency of the first constituent for both.

4.7. Plausibility Effects

Plausibility manipulations have been widely used in the context of studies of sentence parsing. In this section, we will briefly consider the extent to which plausibility/anomaly effects have immediate effects on eye movements. Although there are a few such studies (Braze, Shankweiler, Ni, & Palumbo, 2002; Murray & Rowan, 1998; Ni, Crain, & Shankweiler, 1996; Ni, Fodor, Crain, & Shankweiler, 1998), we will focus on a recent study by Rayner, Warren, Juhasz, and Liversedge (2004) because it has the virtue that the target word was identical across conditions. Rayner et al. had participants read sentences such as

1. John used a knife to chop the large carrots for dinner last night.
2. John used an axe to chop the large carrots for dinner last night.
3. John used a pump to inflate the large carrots for dinner last night.

Sentence 1 is a normal control condition; in sentence 2, it is somewhat implausible that John would use an axe to chop carrots (though one could conceive of such a scenario, as in a camping trip); sentence 3 creates an anomalous scenario. In all sentences, the target word is *carrots*. Rayner et al. found that while sentence 2 only caused mild disruption to reading (and it generally occurred after fixating on the target word), sentence 3 caused immediate disruption, but the disruption occurred rather late in processing the target word (i.e., after the first fixation). Rayner et al. concluded that the default situation in reading is for lexical processes to drive the eyes through the text in reading, but when something does not compute at a higher level (as with the anomalous sentences), then higher-order processes could intervene and cause the eyes to fixate longer (though the influence of such higher-order effects manifest themselves in relatively late processing) time measures on a target word (i.e., in the gaze duration and go-past measure, but not in the first fixation or single-fixation duration measure). What is still not clear, however, is how to objectively define the difference between the implausible and anomalous sentence contexts.

4.8. Languages other than English

In all of our preceding discussions, we have focused largely on results of studies with English speaking readers. However, virtually all of the results that we have discussed hold for other alphabetic writing systems (and indeed some of the studies cited above have been done in other languages). But, we hasten to note that characteristics of the writing system can have powerful influences on eye movements. Here, we mention two such writing systems: Hebrew and Chinese.

Hebrew is a more densely packed language than English, because the vowels are systematically deleted for skilled readers and function words are added as clitics to the end of content words. Of course, Hebrew is printed from right-to-left. These two facts result in the perceptual span being asymmetric to the left of fixation for Hebrew readers (Pollatsek, Bolozky, Well, & Rayner, 1981), as well as smaller than readers of English (for skilled readers of Hebrew, the span extends about 3 letters to the right of fixation to about 11 letters to the left). Interestingly, whereas morphology seems to have little influence on preview benefit for readers of English, readers of Hebrew do show morphological preview benefit (Deutsch, Frost, Peleg, Pollatsek, & Rayner, 2003; Deutsch, Frost, Pollatsek, & Rayner, 2000, 2005). Lexical variables that have been studied in Hebrew seem to yield similar results to English, and Hebrew readers show systematic landing position effects that are similar to English (Deutsch & Rayner, 1999).

Chinese is obviously even more densely packed than Hebrew. This results in a very small perceptual span, which extends from 1 character to the left of fixation to 2–3 characters to the right when reading from left-to-right (Inhoff & Liu, 1997, 1998), and much shorter saccades than English. While the concept of a word is not as well-defined in Chinese as it is in English (and Chinese readers often disagree concerning where word boundaries are located), most words consist of two characters (and most Chinese characters are like morphemes). It has recently been demonstrated that Chinese readers show frequency effects (Yan, Tian, Bai, & Rayner, 2006) and predictability effects (Rayner, Li, Juhasz, & Yan, 2005) that are quite comparable to readers of English.

4.9. Summary of Eye Movements and Word Recognition

Up to this point, we have reviewed some basic findings regarding how certain variables related to word-recognition mechanisms manifest themselves in the eye-movement record. In general, the primary assumption is that lexical factors play a large role in influencing when the eyes move. We do acknowledge that some of the effects we have discussed above are undoubtedly related to post-lexical processing. This raises the question of whether lexical or post-lexical processing, or both, is involved in the decision to move the eyes from one word to the next. Our bias is that many of the effects described above (though obviously not all of them) are reflecting lexical processing. As we will see later, the most successful models of eye-movement control are based on the premise that how long readers look at a word is influenced by the ease or difficulty associated with accessing the meaning of the word. Up to this point, word frequency and word predictability are primarily the indices that have been utilized in the models to predict fixation times. However, in the context of the E–Z Reader model, some effects of morphological complexity (Pollatsek, Reichle, & Rayner, 2003) and number of meanings (Reichle, Pollatsek, & Rayner, 2005b) have been modeled. Our bias is that lexical processing is the engine that drives the eyes and that higher-order effects most likely have influences when something does not compute (as with the anomaly study by Rayner et al., 2004 above). We turn now to the more difficult issues of the effect of higher-order variables (such as parsing and discourse factors) on eye movements.

5. HIGHER-ORDER EFFECTS ON EYE MOVEMENTS: PARSING AND SYNTACTIC AMBIGUITY

Our review of the variables listed in section 4 suggests that a fairly clear, if incomplete, picture is developing with respect to how word processing/lexical factors influence eye movements during reading. However, the same is not true regarding higher-level factors (Clifton et al., 2006). Indeed, effects of parsing/syntactic ambiguity and discourse level variables seem to be highly variable in terms of how they influence eye movements. We will first discuss research on parsing and syntactic ambiguity, and then move to a discussion of the influence of discourse processing on eye movements.

Research on eye movements and syntactic ambiguity resolution has played a central role in the development of theories of sentence processing. It is beyond the scope of this chapter to discuss the extent to which a serial syntax-first type of theory (such as the Garden Path theory presented by Frazier, Clifton, Rayner, and colleagues) or a constraint-satisfaction type of theory (as championed by McDonald, Tanenhaus, and colleagues) can best account for sentence processing. Here, we will focus on the relationship between eye movements and parsing.

Some of the earliest eye-movement research on parsing and syntactic ambiguity held the promise that syntactic factors might have clearly identifiable influences on readers' eye movements. Frazier and Rayner (1982) examined the reading of sentences like (4) and (5), and found that first-fixation durations on the disambiguating region (underlined in the examples) were longer when a temporary ambiguity was resolved in favor of the un-preferred reading (in 4, when *this* was absent). This disruption persisted through the next several fixations, and also appeared as an increased frequency of regressions. Eye movements thus appeared to provide a clear window onto syntactic garden-path effects.

4. Since Jay always jogs a mile and a half (this) <u>seems like</u> a very short distance to him.
5. (The lawyers think) his/the second wife will claim the entire family inheritance (<u>belongs to</u> her).

Much of the disruption in (4) appeared in a region that followed the absence of an obligatory comma (or prosodic break), and disruption in (5) appeared in a sentence-continuation that had no counterpart in the non-disruptive control condition. These facts led to some concerns about this early work. But the force of the missing-comma criticism (i.e., that disruption was caused by the 'mistake' in the materials) is compromised by the fact that an equally obligatory comma was missing in the control condition, with no effect on reading times, and the lack of a closely matched control in (5) was corrected in later research (Rayner & Frazier, 1987).

Frazier and Rayner's results suggested that syntactic processing difficulty could be identified by quickly appearing disruptions in the eye-movement record. Rayner et al. (1983) further provided evidence for a similar conclusion about semantic processing difficulty. They found increased first-pass reading times for the disambiguating region (as well as increased durations of the first three fixations in this region) for sentences like (6), where the first noun is semantically anomalous under the presumably preferred initial analysis, compared to sentences like (7).

6. The kid hit the girl with a wart before he got off the subway.
7. The kid hit the girl with a whip before he got off the subway.

Another early demonstration of syntactic effects on eye movements was presented by Ferreira and Clifton (1986), who showed disruption in the disambiguating region of temporarily ambiguous sentences, both when the initial noun was animate (8) and when it was inanimate (9) and implausible as the subject of the following verb.

8. The defendant (who was) examined by the lawyer proved to be unreliable.
9. The evidence (that was) examined by the lawyer proved to be unreliable.

The disruption appeared in first-pass reading time, and it was argued that the semantic implausibility of the presumably preferred main clause analysis in (9) did not override readers' initial syntactic parsing preferences. However, Trueswell, Tanenhaus, and Garnsey (1994) argued that there were problems with some of Ferreira and Clifton's items and challenged their conclusion. They prepared what they considered to be more adequate sets of materials (which they carefully normed), and found that any effect of ambiguity on first-pass reading time was nonsignificant (indeed, nearly zero, in one experiment) in materials like (9), where semantic preferences weighed against the main clause analysis. They concluded that semantic factors could overturn syntactic preferences, favoring an interactive, constraint-satisfaction, model over the modular serial model favored by Ferreira and Clifton (1986).

Clifton et al.(2003) revisited this issue using materials taken from Trueswell et al. (1994). They varied parafoveal preview of the disambiguating information (since Trueswell et al. made interesting claims about the extent to which readers could use parafoveal information to disambiguate a temporary ambiguity) and participants' reading span. These two manipulations for the most part did not affect the magnitude of the disruption triggered by a temporary ambiguity and the first-pass time measures were similar to those reported by Trueswell et al. (1994). Semantic biases reduced the first-pass reading time measure of the temporary ambiguity effect to non-significance in sentences like (9) (although, similar to Trueswell et al., the interaction of semantic bias and temporary ambiguity was not fully significant, and, unlike Trueswell et al. the ambiguity effect did not go to zero). However, a very different pattern of results was observed for the go-past time and proportion of first-pass regressions out measures. These measures showed disruptive effects of temporary ambiguity that were at least as large in semantically biased inanimate-subject sentences like (9) as in animate-subject sentences like (8) where no semantic bias worked against the presumed preference for a main clause analysis. Clifton et al. (2003) concluded that a full examination of the eye-movement record indicated that initial syntactic parsing preferences were not overcome by semantic biases, although such biases clearly affected overall comprehension difficulty for temporarily ambiguous and unambiguous sentences.

A subsequent analysis of the Clifton et al. (2003) data by Clifton et al. (2006) revealed that while an increase in regressions was responsible for the appearance of a garden-path effect in the inanimate subject condition, regressions were really quite infrequent, always <13% of the trials. This means that the garden-path effects that Clifton et al. (2003) observed in the inanimate-subject condition actually reflected eye-movement events that took place on a minority of the trials. On most trials in the inanimate-subject condition, eye-movements were not affected by temporary ambiguity. It is quite possible that the same holds true for the animate-subject condition: first-pass fixation durations may have been increased by temporary ambiguity on only a small minority of trials. This contrasts sharply with what is true of effects of lexical frequency on fixation durations, where the distribution shifts upwards for low-frequency words (Rayner, 1995; Rayner et al., 2003). No existing research on syntactic garden paths provides data on a large enough number of sentences to permit a convincing distributional analysis to be made (Clifton et al., 2006). It remains a challenge to researchers to devise a way of asking the question of whether first-pass reading times typically or exceptionally increase upon the resolution of a garden path.

In this section so far, we have focused on one difference in the literature on parsing with two studies that utilized the same manipulation, but which came to somewhat different conclusions, depending on which eye-movement measures were focused on. We suspect that this is not an isolated phenomenon (see Binder, Duffy, & Rayner, 2001; Clifton et al., 2006 for further discussion) and that exactly when a given effect will show up in the eye-movement record depends very much on the exact nature of the manipulation and the type of ambiguity present in the study.

In light of the findings we discussed earlier concerning lexical ambiguity resolution, an interesting question is whether the presence of two possible syntactic analyses slows reading, similar to when reading times are slowed when a word has two meanings that are roughly equivalent in frequency? Another question is how are eye movements affected when subsequent material reveals that the reader's initial analysis of a syntactic ambiguity is incorrect. Interestingly, the majority of studies on syntactic ambiguity have not reported any statistically significant effects on reading time in the ambiguous region itself (Staub & Rayner, 2006). A few studies (Frazier & Rayner, 1982; Traxler, Pickering, & Clifton, 1998; van Gompel, Pickering, Pearson, & Liversedge, 2005; van Gompel, Pickering, & Traxler, 2001) have found that an ambiguous region was in fact read more quickly than the corresponding region of an unambiguous control sentence. A few studies have also reported a slowdown in the ambiguous region compared to an unambiguous control (Clifton et al., 2003; Kennison, 2001; Ni et al., 1996; Paterson, Liversedge, & Underwood, 1999; Schmauder & Egan, 1998). However, an explanation other than ambiguity is often available (Staub & Rayner, 2006).

In sum, there is little evidence to indicate that syntactic ambiguity per se causes reading to slow down, and there seem to be circumstances in which ambiguity leads to faster reading times. Evidently, readers either do not consider multiple syntactic analyses in parallel (Frazier, 1978, 1987), or if they do, competition between these analyses does not disrupt processing (van Gompel et al., 2001, 2005). This conclusion stands in contrast with the conclusion from studies of the processing of lexical ambiguity, in which it has been clearly shown that competition between multiple word meanings slows reading.

Eye-movement data have been used to investigate the human parser's preferred analysis of many types of temporary ambiguity; an extensive list of references organized by the type of ambiguity under investigation appears in Clifton et al. (2006). Because there are reliable signs of disruption in the eye-movement record when an initial syntactic analysis is disconfirmed, it has been possible to test subtle and linguistically sophisticated hypotheses about how the parser constructs this initial analysis, and the factors that can influence this analysis. Eye-movement data have helped to reveal the parser's strategies for resolving "long-distance dependencies", in which a phrase appears some distance from the element from which it gets its thematic role, as in the question *Which boy did the teacher reward?*, where *which boy* is the object of the verb *reward* (Pickering & Traxler, 2001, 2003; Traxler & Pickering, 1996). They have also helped to reveal the processing implications of a phrase's status as an argument or adjunct of a verb (Clifton, Speer, & Abney, 1991; Kennison, 2002; Liversedge, Pickering, Branigan, & van Gompel, 1998; Liversedge, Pickering, Clayes, & Branigan, 2003; Speer & Clifton, 1998).

There are open questions about the circumstances under which disambiguation results in a slowing down of forward saccades, regressive eye movements, or both (Altmann, 1994; Altmann, Garnham, & Dennis, 1992; Rayner & Sereno, 1994b, 1994c). However, both Frazier and Rayner (1982) and Meseguer, Carreiras, and Clifton (2002) demonstrated that when readers make regressive eye movements, they do not do so randomly. Instead, where these regressions go reflects some awareness of the point at which the reader's initial, incorrect analysis diverged from the correct analysis.

It is clear that eye-movement data have allowed researchers to probe the early stages of reading in a clear and direct fashion that is exceeded by no other technique. However, a survey reported by Clifton et al. (2006) showed that there was considerable variability in *when* a given manipulation had an effect; this often depended on the type of syntactic construction being used. Clifton et al. noted, and we would certainly agree, that eye-movement data have shown that much, if not quite all, of sentence comprehension is nearly immediate (within a fixation or so after encountering a critical word), as indicated by effects of syntactic or semantic anomaly or complexity and recovery from "garden paths". Eye-movement data have also shown that syntactic knowledge and at least some kinds of semantic, pragmatic, and real-world knowledge have effects even during fixations on the phrase that provides access to this knowledge. But their survey of the literature also clearly showed that the effects of such kinds of knowledge are more variable, even more ephemeral, than the effects that lexical frequency and lexical ambiguity have on eye movements. Fundamental questions, such as whether high-level knowledge consistently affects fixation durations or affects them only now and then, remain unanswered. Furthermore, as Clifton et al. noted, there are disagreements about if and how one kind of knowledge (e.g., knowledge of the situation a sentence describes) modulates the effects of another kind of knowledge (e.g., knowledge of possible syntactic configurations), and disagreements about whether any such modulation is in turn modulated by differences in a reader's abilities and strategies.

In the end then, it is clear that higher-level variables that affect sentence processing and interpretation are much more complex, both in their definition and in their effect, than the variables that govern much of the variation in word recognition. It may be that understanding how these high-level variables operate is not something that can be induced from observations of eye-movement data (as has been true in large part in the domain of word recognition). Rather, as Clifton et al. (2006) noted, understanding must be guided by the development of more explicit theories than now exist of how syntactic, semantic, pragmatic, and real-world knowledge guide language processing.

6. HIGHER-ORDER EFFECTS ON EYE MOVEMENTS: DISCOURSE PROCESSES AND INFERENCES

Whereas eye-movement data are often considered to be the gold standard in studies of sentence parsing, eye-movement data have had much less impact on studies dealing with discourse processes and inferences. In many respects this is quite surprising because it would seem important to determine exactly when readers make inferences as they read, and certainly eye-movement data hold the promise of revealing this type of temporal effect. In this section, we will review studies that have used eye movements in this manner. Our suspicion is that proportionally more eye-movement studies dealing with discourse processes and inferences will appear over the next few years.

In understanding text, readers must be able to integrate information within sentences and also make connections across sentences to form a coherent discourse representation. To what extent can eye-movement data reflect these processes? In this section, we will review research dealing with sentence and clause-wrap up, antecedent search, and on-line inferences.

6.1. Sentence and Clause Wrap-up

Just and Carpenter (1980) found that fixation times on words that occurred at the end of a sentence were unusually long (in comparison to words that did not end a sentence) as measured by a regression analysis. Subsequently, Rayner et al. (1989) reported that when a target word ended a clause or a sentence, fixation times were inflated in comparison to when that same word did not end a clause or sentence. More recently, Rayner, Kambe, and Duffy (2000) confirmed this finding and further demonstrated that not only were fixations longer on clause and sentence final words, but that the next saccade was lengthened (see also Hill & Murray, 2000). So, readers slow down at clause and sentence boundaries, but then send their eyes further into the next region of text as if processing capacity had been freed up once the wrap-up processes had been completed. Hirotani et al. (2006) have followed up on these findings and demonstrated that implicit prosody/intonation (imposed internally by the reader) is also very much involved in wrap-up effects.

6.2. Antecedent Search

The process of establishing a connection between an anaphoric element (such as a pronoun) and its antecedent in the text, antecedent search, is central to comprehending discourse. Pronominal reference and noun–noun reference are two such instances in which the correct linkage between discourse elements is required for text comprehension.

In pronominal reference, when a pronoun like *sh*e is encountered in the course of reading, the reader must identify an antecedent that matches it in number and gender. Sometimes, the process is trivially easy and no disruption is observed in the eye-movement record (Blanchard, 1987). If the pronoun involves violation of a gender stereotype (referring to a truck driver as *she*), fixations are inflated (Duffy & Keir, 2004; Sturt, 2003). If there is considerable distance between the pronoun (or anaphor) and the antecedent, readers' fixations are longer when the pronoun is encoded and the antecedent search may continue over the next couple of fixations (Ehrlich & Rayner, 1983; Garrod, Freudenthal, & Boyle, 1994; O'Brien, Raney, Albrecht, & Rayner, 1997); if the distance between pronoun/anaphor and antecedent is close, fixations are not inflated as much. Of course, since pronouns are typically short words, readers skip over them quite frequently (thus making it difficult to determine exactly when the pronoun was encoded). Interestingly, van Gompel and Majid (2004) found that pronouns with infrequent antecedents yielded longer fixations in the encoding region than pronouns with more frequent antecedents; the effect did not occur on the pronoun itself but was slightly delayed to the region following the pronoun.

Just as a pronoun requires an antecedent, a definite Noun Phrase (NP) that does not directly refer to something outside the text requires a coreferring antecedent in the text. Thus, if a reader encounters the NP *the bird* after earlier reading about a robin (or vice versa), there is an antecedent link. Whereas pronouns carry little semantic information beyond gender and number, nouns typically have more semantic content, which facilitates the search for the antecedent. Duffy and Rayner (1990) found evidence that antecedent search time with anaphoric NPs was primarily localized on the target noun (see also van Gompel & Majid, 2004), so that there were no major spillover effects as with pronouns.

6.3. On-line Inferences

Within a discourse representation, the simplest kind of connection is one in which one word gains its reference through another word in the text, such as anaphor. However, elaborative inferences occur when information that has not been explicitly stated up to a given point in the text is inferred by the reader. Eye-movement data have confirmed that elaborative inferences occur on-line during reading. These data have also served to differentiate between conditions in which readers make an inference and those in which they wait for more explicit information.

O'Brien, Shank, Myers, and Rayner (1988) asked readers to read short passages of text and fixation time was examined on a target word (*knife*) in the final sentence of a passage as in (10).

10. He threw the knife into the bushes, took her money, and ran away.

The target word was previously either explicitly mentioned in the text (as in the phrase *he stabbed her with his knife*) or only strongly suggested (as in *he stabbed her with his weapon*). O'Brien et al. (1988) found no difference in gaze duration on the target word across these two conditions. It thus appears that the concept knife had been inferred from the prior context when the word *weapon* was actually present. In contrast, when the text did not strongly suggest the concept of a knife (as in *he assaulted her with his weapon*), gaze duration on the target word *knife* was longer compared to when it had been explicitly mentioned or strongly suggested earlier in the passage. These results indicate that the longer fixation time on the target word was due to a memory search for its antecedent and that antecedent search begins immediately upon fixating the target word.

Although O'Brien et al. (1988) found evidence for on-line elaborative inferences, they also demonstrated that such inferences only occur when there is a "demand sentence" (which invited the reader to make the inference) just prior to the sentence containing the target word. A subsequent study by Garrod, O'Brien, Morris, and Rayner (1990) further constrained the conditions under which elaborative inferences occur. Their data suggested that the presence of a demand sentence invites the reader to actively predict a subsequent expression and elaborative inferences only occur when there is an anaphoric relationship between two nouns.

Other studies have observed rather immediate effects in the eye movement record of bridging inferences (Myers, Cook, Kambe, Mason, & O'Brien, 2000) and the integration of role fillers in scripted narratives (Cook & Myers, 2004; Garrod & Terras, 2000). The most interesting point in these studies is that higher-order variables show immediate effects in the eye-movement record. At some level, studies such as these provide a problem for models of eye-movement control in reading, which are largely based on the assumption that lexical processing is the engine that drives the eyes through the text in reading. We now turn to a discussion of such models.

7. MODELING EYE MOVEMENTS IN READING

Obviously, developing a quantitative model that could explain all the phenomena that have been observed in reading is a task that is beyond us at present, and it may be an unreachable goal. That is, there are so many factors that influence reading, ranging from the legibility of the characters up through the frequency of the words in the language, the complexity of the syntax, the higher-order organization of the text, and the real-world knowledge shared by the author and the reader. If a model were to try to explain all of these factors, it would almost certainly either be hideously complex or degenerate into a multiple regression equation that merely re-affirmed that all of these variables (and others) play a part in reading.

As a result, our belief is that for a model of reading to be of some value at this point in time, it needs to be able to explain a significant part of the reading process, yet be simple enough so that it is a useful heuristic tool for understanding which phenomena it can explain and which it can not explain. At present, there are a number of programs of research that are developing quantitative models of eye movements in reading. Though they differ wildly in many respects, they all share certain features, as none are attempting any serious modeling of how text is parsed or how discourse structures are being constructed. Some focus on lexical processes primarily influencing eye movements, whereas others attempt to explain eye movements in reading largely by lower-level, perceptual and motor processing or via Ideal Observer procedures. The primary models are E–Z Reader (Pollatsek, Reichle, & Rayner, 2006; Rayner, Ashby, Pollatsek, & Reichle, 2004; Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle, Pollatsek, & Rayner, 2006b; Reichle, Rayner, & Pollatsek, 1999, 2003), SWIFT (Engbert, Longtin, & Kliegl, 2002; Engbert, Nuthmann, Richter, & Kliegl, 2005; Kliegl & Engbert, 2003), Glenmore (Reilly & Radach, 2003), SERIF (McDonald, Carpenter, & Shillcock, 2005), Mr. Chips (Legge, Klitz, & Tjan, 1997; Legge, Hooven, Klitz, Mansfield, & Tjan, 2002), and the Competition/Interaction model (Yang & McConkie, 2001, 2004). These models differ on a number of dimensions, but space does not permit us to discuss the models in detail (though most of them were reviewed by Reichle et al., 2003). As a result, we will largely try to illustrate this modeling enterprise through outlining the modeling enterprise we are associated with (versions of the E–Z Reader model), and will briefly comment on how some of the other models differ and the points of controversy.

7.1. The E–Z Reader Model

The E–Z Reader model focuses on trying to explain how lexical processing influences the progress of the eyes through the text. We think this is a justifiable focus, as a case can be made that higher-order variables (such as constructing discourse structures) have a more indirect influence on how readers make their way through the text. That is, individual fixation durations on words are typically about 250 ms and gaze durations typically average no more than 300–350 ms, and the motor programming time for an eye movement is far from instantaneous, as the time necessary to make an eye movement to the simplest visual stimuli (e.g., the onset of a point of light) takes close to 200 ms. Thus, it seems unlikely that readers are waiting for all levels of processing to be completed (e.g., constructing a parsing tree of the sentence to that point in the text) before sending a signal to the eye-movement system to move on to the next word.

Instead, we think it is a reasonable working hypothesis that linguistic processing affects eye movements in two different ways. First, there is a relatively low-level of linguistic processing that keeps the eyes moving forward – we have tentatively associated this with lexical processing although this might be too restrictive. Second, higher-level processing is occurring in parallel with this lexical processing system, and when it becomes clear that the higher-level processing is having difficulty, either because the higher-order processing system is falling behind the encoding of words or because an "error" has been detected (such as occurs when readers read something that is anomalous or when they misparse a sentence), this second system intervenes to tell the first system to either stay in place until the second system catches up or to go back and make an attempt to repair the damage. We think that such a hypothesis makes sense, as reading would proceed far more slowly than the normal 300 words per minute if the reader had to wait on each word until its significance within the text was ascertained. Whether this is in fact how reading goes on is, of course, an open question. However, we think it is a reasonable starting point for thinking about the reading process. In addition, modeling of the first stage then no longer seems like an insurmountable task; moreover, it allows one to define a well-defined set of eye-movement data to be modeled: all eye movements in reading other than regressions back to prior words. Let us see how this might be done.

In the E–Z Reader model, we have posited two more-or-less modular systems: the cognitive system and the motor system. Thus, there are sets of assumptions relating to the events in the cognitive system that trigger eye movements and sets of assumptions about how the commands to execute eye movements actually get carried out. In E–Z Reader, the first basic assumption is that a stage of lexical access causes a program for an eye-movement to the next word in the text to be initiated, and the second basic assumption is that this eye-movement program is executed within about 150 ms after it is initiated – in the normal state of affairs. Obviously, the above two assumptions can not be the whole story as they would predict that each word is fixated exactly one time, and we know that some words are skipped and that others are fixated more than once. Thus, other assumptions need to be made as well. In addition, one needs to include some model of covert attention in a model; that is, one has to make assumptions about what is being processed at any moment in time. It is the type of attentional assumption that is perhaps the major distinguishing feature of the various reading models. In E–Z Reader, it is assumed that low-level visual processing goes on in parallel over the whole visual field – such low-level processing, among other things, allows the eye-movement system to be able to target saccades. In contrast, E–Z Reader assumes that lexical processing is serial in the sense that only one word is being processed at any moment in time. However, we want to emphasize that this does not mean that only one word is processed on a fixation; on the contrary, the usual state of affairs in the E–Z Reader model is that two words are processed on a fixation, and, not infrequently, at least partial processing of three words occurs on a fixation.

As just indicated, the key assumption is that words are attended to (and thus lexically processed) one at a time. The simplest possible assumption about how this attention management would occur was made by Morrison (1984); he posited that when the reader lexically accesses a word (a) an eye-movement program is initiated and (b) attention shifts to the next word. For reasons we discuss later, we thought this was too simple, so we chose a slightly more complex model. We assumed that an earlier stage of lexical access (L1) is the trigger for an eye movement, but that a later stage of lexical access (L2) is the trigger for an attention shift and hence the start of processing the next word. One can view this system as one in which readers have developed a "cheat" and trigger an eye movement (which takes appreciable time to execute) when they are reasonably sure that the word has been comprehended (L1), but only start to process the next word after the lexical process has completed (L2). In the model, we assume that both (a) the duration of the first stage and (b) the time between the completion of the first stage and the completion of the second stage are linear functions of log frequency. (We will subsequently refer to the difference in time between when L1 and L2 are completed as the duration of the L2 stage.) We also assume that the durations of both stages are affected by the predictability of the word from the prior text. In earlier versions of the model, we assumed that the influence of predictability was multiplicatively related to the influence of word frequency, but then realized (Rayner, Ashby et al., 2004) that an additive version (i.e., that frequency and predictability made independent contributions to the speed of lexical access) was better.

There is one more assumption about how eye movements are triggered. This is an assumption related to refixating the currently attended word. In an earlier version, we assumed that a refixation on the currently attended word was programmed automatically when a fixation began. (We will discuss below why this does not imply that all words are refixated.) However, there were problems with this simple mechanism for refixations on words, and in the current version, refixations are programmed (a) with a probability <1 when a word is fixated and (b) the probability depends on how far from the center of the word the fixation is. However, as the refixation component is not a particularly well-worked out aspect of the model, we will only give a sketch how these mechanisms can explain refixations.

Now we turn to the assumptions about the programming and execution of eye movements. The key assumption (adapted from Morrison, 1984) is that later eye-movement programs *can* cancel earlier eye-movement programs. This assumption is based on the work of Becker and Jürgens (1979), who examined a much simpler situation than reading. They had participants fixate a small area of light at point 1, which then moved abruptly to point 2. When this was all that happened, people quickly fixated point 2. The key trials were when the light moved abruptly again to point 3. If the gap in time between the two movements was sufficiently large, participants fixated point 2 and then point 3. However, when the gap was short enough, they merely fixated point 3, indicating that they could cancel the eye movement to point 2.[3] To capture saccade cancellation in our modeling, we assume that there are two stages in a motor program: a labile stage followed by a non-labile stage. If an

[3]At intermediate times, sometimes there were compromise saccades. We have modeled such compromise saccades and there was little change in the predicted values in the simulations.

eye movement program is in the labile stage, a subsequent eye-movement program can cancel it, whereas if it is in the non-labile stage, it will be executed regardless of what other programs are initiated. (In the latest versions of our model, the former stage is assumed to be about 100 ms and the latter stage about 50 ms.)

Let us see how this relates to reading. First, let us examine skipping. What words are skipped? Mainly words that take little time to process such as frequent and/or predictable words. In E–Z Reader, as we shall see, a reader typically has done some processing of a word before it is fixated and then finishes processing the word when it is fixated and the signal from the completion of the L1 stage typically occurs about 100–150 ms after the word is fixated. This produces a signal to fixate the next word in the text (word $n+1$). Not that much after this (the L2 duration, which is typically $\sim$50 ms), attention moves to word $n+1$ and lexical access of it begins. However, if word $n+1$ is easy to process, stage L1 will be quick, and can finish before the end of the labile stage of the eye-movement program to fixate word $n+1$. The completion of the L1 stage of word $n+1$ in these cases thus produces a program to fixate word $n+2$ which will then cancel the program to fixate word $n+1$, and hence word $n+1$ will be skipped. Moreover, as we have argued above, the model predicts that the probability that this will happen will be greater for more frequent and/or more predictable words.

Before moving on, let us briefly sketch how this cancellation assumption affects refixations. This is easier to do if we use our earlier simple assumption that a program to refixate a word is made automatically upon first fixating a word. As we indicated above, this assumption would cause all words to be refixated unless the refixation program is cancelled. When will the refixation program be cancelled? Answer: when the program to fixate word $n+1$ occurs during the labile stage of this refixation program. This will occur when word n (the fixated word) is easy to process because it is high frequency and/or predictable. Thus, the model predicts that lower-frequency words and less predictable words are more likely to be refixated. (We should emphasize that we assume that the eye-movement system knows nothing about cognition; thus all eye-movement programs are assumed to have the same properties, regardless of the triggering mechanism.)

In addition, one needs to make assumptions about the targeting of the saccades. There are two issues involved here. The first is to specify exactly what the target of a saccade is. The second is to posit the error involved in the targeting procedure. In E–Z Reader, for both of these issues, we basically imported the data and assumptions from work by McConkie et al. (1998). We assume that the target of a saccade is the center of a word, but that this is subject to both a constant error (i.e., short saccades, on average, will tend to overshoot the target location and long saccades, on average, will tend to undershoot the target location) and random error. These assumptions give a pretty good account of the landing positions on a word. We should also emphasize that they also imply that the target word (i.e., the attended word) is not always the fixated word (due to noise in the oculomtor system). One more processing assumption is worth mentioning in this regard. That is, that the speed of lexical processing does not only depend on the frequency of a word and its predictability, but also where its letters are with respect to the fixation point – the further these letters are from fixation, the slower processing is (due to visual acuity concerns). This implies not only that processing a word in the parafovea before it is fixated is less efficient than when it is fixated, but that longer words will be processed more slowly than shorter words (all else being equal) because, on average, the letters of longer words will be further from fixation. Moreover, this length effect is magnified, the further the fixation point is from the center of the word.

This, in outline, summarizes the E–Z Reader model. As we hope we have indicated, it can account qualitatively for many of the major phenomena of reading related to word variables. It predicts that longer, less frequent, and/or less predictable words will be (a) fixated longer, (b) skipped less, and (c) refixated more often. Moreover, it gives a very good quantitative account of these phenomena for sentence reading (see Pollatsek et al., 2006; Rayner et al., 2003). Furthermore, it does so using quite reasonable assumptions about how long lexical processing takes and how long motor programs take. We make the latter point, because there is still some skepticism that cognitive processes in reading can possibly be fast enough to influence eye movements in an on-line fashion.

Before going on to briefly discuss the competing models, we need to return to a point that we quickly slid over before: the motivation for positing two stages of lexical processing rather than just have a single stage be the trigger for an eye-movement program and an attention shift. One reason is that, if one assumes that there is a single stage of lexical processing, there can be no delayed effects due to difficulty in lexical processing. That is, a one-stage model would predict that one continues to process (and fixate) a word until it is processed, and then, simultaneously, (a) an eye movement is programmed to fixate word $n+1$ and (b) attention shifts to word $n+1$. Thus, the time that word $n+1$ will be processed in the parafovea before it is fixated will be merely equal to the latency of the eye-movement program and will not be a function of the difficulty of word n. However, as indicated in an earlier section of this chapter, there are many findings that the difficulty of word n often "spills over" to affect the time taken to process word $n+1$. Our assumption that the duration of the second stage of word processing is also a function of difficulty of lexical processing explains such spillover effects. We should point out, however, that the E–Z Reader model does not predict such spillover effects when early stages of word identification are manipulated. In one such manipulation (Reingold & Rayner, 2006), a target word was made quite faint and this increased the gaze durations on this word by over 100 ms. However, in this case there were no spillover effects, consistent with the likelihood that this manipulation only affected early stages of lexical processing.

A second motivation for our assumption of two stages of lexical processing is that it nicely explains the phenomenon that there is less preview benefit when the fixated word is more difficult to process (e.g., Henderson & Ferreira, 1990; Kennison & Clifton, 1995; White, Rayner, & Liversedge, 2005). This again, is explained by our assumption that the duration of the second stage of lexical processing is a function of the difficulty of processing the fixated word. (Roughly speaking, in the E–Z Reader model, the amount of time a word is processed in the parafovea is equal to the eye movement latency minus the duration of this second lexical processing stage.) A third reason for our positing two processing stages is that it also gives a significantly better account of the general data. That is, a model that assumes only one stage can actually produce a decent global fit to the reading data (i.e., predict mean gaze durations, mean first-fixation durations, mean skipping rates, and other indices of reading as a function of word frequency, length, and predictability). However, in order to do so, a one-stage model needs to posit that processing of parafoveal material is much more efficient than it is, which, in turn, leads it to predict far too big effects of *preview benefit* (see Section 4 above).

As a result, we feel reasonably comfortable in the claim that our model gives a good overall account of eye movements in reading as long as there are no "higher-order" difficulties in the text. However, it is admittedly vague on some points, such as what the two stages of lexical access are, and there are some conceptual problems that we need to address in the future. However, we will discuss those after briefly discussing some of the other quantitative models of reading, perhaps not from the most unbiased perspective.

7.2. Other Models

One model we will spend little time on is the Competition/Interaction model of Yang and McConkie (2001, 2004). The reason for this is that the major way that their model differs from E–Z Reader is that they posit that lexical processing plays only a minor role in reading. This inference is mainly drawn from a rather unnatural paradigm in which text keeps on disappearing rapidly and which may have little to do with normal reading. Clearly, such a model is at variance with much of the data reviewed earlier, which shows that many linguistic variables play a role in how long words are fixated. They admit that such linguistic variables can play a role, but only in some cases when processing is lengthened appreciably. However, this is at odds with the data that show, for example, that the distribution of fixation durations on less frequent words is essentially the same shape as that for more frequent words, except that the whole distribution is shifted to the right (see Rayner, 1995; Rayner, Liversedge et al., 2003). The variance for the less frequent word, admittedly, is slightly bigger; however, the pattern is not what would be predicted if a large majority of fixations durations were determined by low-level processing and a few long fixations were due to difficulty in processing the low-frequency words.

In contrast, lexical processing is at the core of two of the other competing models (SWIFT and Glenmore). A major difference between them and E–Z Reader is that lexical processing is assumed to go on in parallel in both models. In SWIFT, it is assumed that lexical processing is simultaneously occurring on four words (from the one to the left of fixation up to the two words to the right), whereas in Glenmore, processing only occurs on the fixated word and the one to the right. Moreover, in SWIFT, the assumptions about the relation between lexical processing and eye-movement control are complex; for example, lexical excitation is assumed to have an inverted U-shaped function as far as how the word attracts attention and thus an eye movement. Both models appear to make these parallel assumptions about lexical processing because of so-called parafoveal-on-foveal effects (Kennedy, 2000; Kennedy & Pynte, 2005), in which the duration of fixations on word n are influenced by the properties of word $n+1$. That is, such effects seem incompatible with a serial processing model such as E–Z Reader. Space does not permit a full discussion of this issue, so we will briefly address three points.

First, lexical parafoveal-on-foveal effects are far from an established empirical phenomenon (see Rayner & Juhasz, 2004; Rayner, White, Kambe, Miller, & Liversedge, 2003). That is, although there appear to be reliable parafoveal-on-foveal effects when word $n+1$ is visually or orthographically unusual, parafoveal-on-foveal effects caused by lexical properties are far from established. Thus, at present, there is nothing in this phenomenon that is particularly suggestive that lexical processing is occurring in parallel. Second, E–Z Reader in fact can explain some parafoveal-on-foveal effects, as it predicts that words are occasionally mistargeted; hence, with some reasonable probability the reader intended to fixated word $n+1$ (and is attending to it) but the saccade fell short and the reader is fixating word n (see Rayner, Warren et al., 2004). However, because of the uncertainty of many of the phenomena, we have not tried to model parafoveal-on-foveal effects, so we can not be sure that the model does in fact adequately explain such phenomena. Third, we hesitated to posit parallel processing of words because it does not seem all that psychologically plausible. That is, as current modeling of word identification is having a very hard time in explaining how a single word is identified, we think it is almost like positing magic to say that more than one word is processed in parallel. (There might be exceptions for frequent combinations of short words like “to the”.) More generally, from a modeling standpoint, we think that the serial processing assumption makes the model more transparent, and thus makes it a more valuable heuristic device for understanding reading, and that one should abandon it only if there are compelling reasons to do so.

7.3. Models’ Summary

Let us briefly close this section by commenting on the problems with the model and what we see as what needs to be accomplished in the next 3–5 years in modeling. First, E–Z Reader, like the other models, does not contain a serious model of word recognition. We have merely posited processes that have finishing times influenced by variables such as frequency. This needs addressing. Second, we have been quite vague about what L1 and L2 are, and have merely used vague terms such as “lexical access”, without specifying what codes are accessed (e.g., orthographic, phonological, semantic, syntactic). One arena that this lack is particularly apparent is that, at present, we have no complete satisfactory way of explaining the effects of lexical ambiguity discussed earlier. One can possibly push these effects off to the “higher-order processing system”, but this seems unsatisfactory. We are currently working on this (Reichle et al., 2006a). Second, as indicated earlier, there are certain syntactic and semantic anomalies that have quite immediate effects on eye movements. This suggests that, if one wants to model all immediate eye-movement effects (presumably the scope of the E–Z Reader model and competing models), one has to come up with a satisfactory explanation of why these effects (but not others) are immediate. We think the solution to this problem will be non-trivial. However, let us close on a more positive note. That is, our modeling was previously also fairly vague about its assumptions about early (pre-lexical) visual processing and its relation to lexical processing. We have addressed most of these issues in a recent article (Pollatsek et al., 2006).

8. GENERAL SUMMARY

In this chapter, we have reviewed data on eye movements in reading. It is clearly the case that the ease or difficulty associated with the fixated word strongly influences how long the reader will look at that word. Current models of eye-movement control have seized heavily on this fact and have been very successful in accounting for a large amount of the variance in how readers move their eyes through text. But, as we have seen, there are also clearly other influences on eye movements besides low-level lexical factors. Nevertheless, it does appear that a great deal can be accounted for via the assumption that lexical processing (or the ease or difficulty associated with the fixated word) is the engine driving the eyes through the text and that higher-order information primarily serves to intervene when something does not compute well.

Finally, we hope it is obvious that eye movements continue to be an excellent way to study the moment-to-moment processes inherent in the reading process. As we have seen in our discussion of higher-order factors, there are many difficulties in interpreting the eye-movement record when the manipulations involve syntax or higher-order semantic processing. However, it terms of inferring moment-to-moment mental processing during reading, it is not clear that there are any other measures that provide the temporal precision that eye movements provide.

ACKNOWLEDGMENT

Preparation of this chapter was made possible by grants HD17246 and HD26765 from the National Institute of Health.

REFERENCES

Altarriba, J., Kambe, G., Pollatsek, A., & Rayner, K. (2001). Semantic codes are not used in integrating information across eye fixations in reading: Evidence from fluent Spanish–English bilinguals. *Perception & Psychophysics*, *63*, 875–890.

Altarriba, J., Kroll, J. F., Sholl, A., & Rayner, K. (1996). The influence of lexical and conceptual constraints on reading mixed language sentences: Evidence from eye fixation and naming times. *Memory & Cognition*, *24*, 477–492.

Altmann, G. T. M. (1994). Regression-contingent analyses of eye movements during sentence processing: A reply to Rayner and Sereno. *Memory & Cognition*, *22*, 286–290.

Altmann, G. T. M., Garnham, A., & Dennis, Y. (1992). Avoiding the garden path: Eye movements in context. *Journal of Memory and Language*, *31*, 685–712.

Andrews, S., Miller, B., & Rayner, K. (2004). Eye movements and morphological segmentation of compound words: There is a mouse in mousetrap. *European Journal of Cognitive Psychology*, *16*, 285–311.

Ashby, J., & Clifton, C. (2005). The prosodic property of lexical stress affects eye movements during silent reading. *Cognition*, *96*, B89–B100.

Ashby, J., & Rayner, K. (2004). Representing syllable information during silent reading: Evidence from eye movements. *Language and Cognitive Processes*, *19*, 391–426.

Ashby, J., Rayner, K., & Clifton, C. (2005). Eye movements of highly skilled and average readers: Differential effects of frequency and predictability. *Quarterly Journal of Experimental Psychology*, *58A*, 1065–1086.

Balota, D. A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology*, *17*, 364–388.

Becker, W. & Jürgens, R. (1979). An analysis of the saccadic system by means of double step stimuli. *Vision Research*, *19*, 967–983.

Bertram, R., & Hyönä, J. (2003). The length of a complex word modifies the role of morphological structure: Evidence from eye movements when reading short and long Finnish compounds. *Journal of Memory and Language*, *48*, 615–634.

Binder, K., Duffy, S., & Rayner, K. (2001). The effects of thematic fit and discourse context on syntactic ambiguity resolution. *Journal of Memory and Language*, *44*, 297–324.

Binder, K. S. (2003). Sentential and discourse topic effects on lexical ambiguity processing: An eye-movement examination. *Memory & Cognition*, *31*, 690–702.

Binder, K. S., & Morris, R. K. (1995). Eye movements and lexical ambiguity resolution: Effects of prior encounter and discourse topic. *Journal of Experimental Psychology*: *Learning, Memory, and Cognition*, *21*, 1186–1196.

Binder, K. S., Pollatsek, A., & Rayner, K. (1999). Extraction of information to the left of the fixated word in reading. *Journal of Experimental Psychology*: *Human Perception and Performance*, *25*, 1162–1172.

Binder, K. S., & Rayner, K. (1998). Contextual strength does not modulate the subordinate bias effect: Evidence from eye fixations and self-paced reading. *Psychonomic Bulletin & Review*, *5*, 271–276.

Birch, S., & Rayner, K. (1997). Linguistic focus affects eye movements during reading. *Memory & Cognition*, 25, 653–660.

Blanchard, H. E. (1987). Pronoun processing during fixations: Effects on the time course of information utilization. *Bulletin of the Psychonomic Society*, *25*, 171–174.

Braze, D., Shankweiler, D., Ni, W., & Palumbo, L. C. (2002). Readers' eye movements distinguish anomalies of form and content. *Journal of Psycholinguistic Research*, *31*, 25–44.

Briihl, D., & Inhoff, A.W. (1995). Integrating information across fixations in reading: The use of orthographic bodies and of exterior letters. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 55–67.

Brysbaert, M., & Vitu, F. (1998). Word skipping: Implications for theories of eye movement control in reading. In: G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 125–148). Oxford: Elsevier.

Calvo, M. G., & Meseguer, E. (2002). Eye movements and processing stages in reading: Relative contribution of visual, lexical, and contextual factors. *Spanish Journal of Psychology*, *5*, 66–77.

Carpenter, P. A., & Daneman, M. (1981). Lexical retrieval and error recovery in reading: A model based on eye fixations. *Journal of Verbal Learning and Verbal Behavior*, *28*, 138–160.

Carpenter, P. A., & Just, M. A. (1983). What your eyes do while your mind is reading. In: K. Rayner (Ed), *Eye movements in reading: Perceptual and language processes* (pp. 275–307). New York: Academic Press.

Carroll, P. J. & Slowiaczek, M. L. (1986). Constraints on semantic priming in reading: A fixation time analysis. *Memory & Cognition*, *14*, 509–522.

Chace, H. E., Rayner, K., & Well, A. D. (2005). Eye movements and phonological parafoveal preview benefit: Effects of reading skill. *Canadian Journal of Experimental Psychology*, *59*, 209–217.

Chaffin, R., Morris, R. K., & Seely, R. E. (2001). Learning new word meanings from context: A study of eye movements. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *27*, 225–235.

Clifton, C. Jr., Speer, S., & Abney, S. (1991). Parsing arguments: Phrase structure and argument structure as determinants of initial parsing decisions. *Journal of Memory and Language*, *30*, 251–271.

Clifton, C. Jr., Staub, A., & Rayner, K. (2006). Eye movements in reading words and sentences. In: R. P. G. von Gompel, M. H. Fischer, W. S. Murray, R. C. Hill (Eds), *Eye movements: A window on mind and brain*. Oxford: Elsevier.

Clifton, C., Traxler, M., Mohamed, M. T., Williams, R. S., Morris, R. K., & Rayner, K. (2003). The use of thematic role information in parsing: Syntactic processing autonomy revisited. *Journal of Memory and Language*, *49*, 317–334.

Cook, A. E., & Myers, J. L. (2004). Processing discourse roles in scripted narratives: The influences of context and world knowledge. *Journal of Memory and Language*, *50*, 268–288.

Daneman, M., & Reingold, E. (1993). What eye fixations tell us about phonological recoding during reading. *Canadian Journal of Experimental Psychology*, *47*, 153–178.

Daneman, M., Reingold, E., & Davidson, M. (1995). Time course of phonological activation during reading: Evidence from eye fixations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 884–898.

Deutsch, A., Frost, R., Peleg, S., Pollatsek, A., & Rayner, K. (2003). Early morphological effects in reading: Evidence from parafoveal preview benefit in Hebrew. *Psychonomic Bulletin & Review*, *10*, 415–422.

Deutsch, A., Frost, R., Pollatsek, A., & Rayner, K. (2000). Early morphological effects in word recognition: Evidence from parafoveal preview benefit. *Language and Cognitive Processes*, *15*, 487–506.

Deutsch, A., Frost, R., Pollatsek, A., & Rayner, K. (2005). Morphological parafoveal preview benefit effects in reading: Evidence from Hebrew. *Language and Cognitive Processes*, *20*, 341–371.

Deutsch, A., & Rayner, K. (1999). Initial fixation location effects in reading Hebrew words. *Language and Cognitive Processes*, *14*, 393–421.

Dopkins, S., Morris, R. K., & Rayner, K. (1992). Lexical ambiguity and eye fixations in reading: A test of competing models of lexical ambiguity resolution. *Journal of Memory and Language*, *31*, 461–477.

Drieghe, D., Rayner, K., & Pollatsek, A. (2005). Eye movements and word skipping during reading revisited. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 954–969.

Duffy, S. A., & Keir, J. A. (2004). Violating stereotypes: Eye movements and comprehension processes when text conflicts with world knowledge. *Memory & Cognition*, *32*, 551–559.

Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language*, *27*, 429–446.

Duffy, S. A., & Rayner, K. (1990). Eye movements and anaphor resolution: Effects of antecedent typicality and distance. *Language and Speech*, *33*, 103–119.

Ehrlich, K., & Rayner, K. (1983). Pronoun assignment and semantic integration during reading: Eye movements and immediacy of processing. *Journal of Verbal Learning and Verbal Behavior*, *22*, 75–87.

Ehrlich, S. F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior*, *20*, 641–655.

Engbert, R., Longtin, A., & Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research*, *42*, 621–636.

Engbert, R., Nuthmann, A., Richter, E., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review*, *112*, 777–813.

Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory & Language*, *25*, 348–368.

Folk, J. R. (1999). Phonological codes are used to access the lexicon during silent reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 892–906.

Folk, J. R., & Morris, R. K. (1995). Multiple lexical codes in reading: Evidence from eye movements, naming time, and oral reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1412–1429.

Folk, J. R., & Morris, R. K. (2003). Effects of syntactic category assignment on lexical ambiguity resolution in reading: An eye movement analysis. *Memory & Cognition*, *31*, 87–99.

Frazier, L. (1978). *On comprehending sentences: Syntactic parsing strategies*. Doctoral dissertation, University of Connecticut, Storrs.

Frazier, L. (1987). Sentence processing: A tutorial review. In: M. Coltheart (Ed.), *Attention and performance XII* (pp. 559–586). Hillsdale, NJ: Lawrence Earlbaum Associates, Inc.

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

Frazier, L., & Rayner, K. (1987). Resolution of syntactic category ambiguities: Eye movments in parsing lexically ambiguous sentences. *Journal of Memory and Language*, *26*, 505–526.

Frazier, L., & Rayner, K. (1990). Taking on semantic commitments: Processing multiple meanings vs. multiple senses. *Journal of Memory and Language*, *29*, 181–200.

Frisson, S., & Frazier, L. (2005). Carving up word meaning: Portioning and grinding. *Journal of Memory and Language*, 53, 277–291.

Frisson, S., Niswander-Klement, E., & Pollatsek, A. (2006). The role of semantic transparency in the processing of English compound words, submitted.

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

Frisson, S., Rayner, K., & Pickering, M. J. (2005). Effects of contextual predictability and transitional probability on eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 862–877.

Garrod, S., Freudenthal, S., & Boyle, E. (1994). The role of different types of anaphor in the on-line resolution of sentences in a discourse. *Journal of Memory and Language*, *33*, 39–68.

Garrod, S., O'Brien, E. J., Morris, R. K., & Rayner, K. (1990). Elaborative inferencing as an active or passive process. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 250–257.

Garrod, S., & Terras, M. (2000). The contribution of lexical and situational knowledge to resolving discourse roles: Bonding and resolution. *Journal of Memory and Language*, *42*, 526–544.

Gautier, V., O'Regan, J. K., & LaGargasson, J.F. (2000). 'The skipping' revisited in French: Programming saccades to skip the article 'les'. *Vision Research*, *40*, 2517–2531.

Haberlandt, K. (1994). Methods in reading research. In: M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp.1–32). San Diego: Academic Press.

Heller, D., & Radach, R. (1999). Eye movements in reading: Are two eyes better than one? In: W. Becker, H. Deubel, & T. Mergner, (Eds) *Current oculomotor research: Physiological and psychological aspects* New York: Plenum Press.

Henderson, J. M., Dixon, P., Petersen, A., Twiley, L. C., & Ferreira, F. (1995). Evidence for the use of phonological representations during transaccadic word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 82–97.

Henderson, J. M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 417–429.

Henderson, J. M., & Ferreira, F. (1993). Eye movement control during reading: Fixation measures reflect foveal but not parafoveal processing difficulty. *Canadian Journal of Experimental Psychology*, *47*, 201–221.

Hill, R. L., & Murray, W. S. (2000). Commas and space: Effects of punctuation on eye movements and sentence parsing. In: A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds), *Reading as a perceptual process* (pp. 565–590). Oxford: Elsevier.

Hirotani, M., Frazier, L., & Rayner, K. (2006). Punctuation and intonation effects on clause and sentence wrap–up: Evidence from eye movements. *Journal of Memory and Language*, *54*, 425–443.

Hyönä, J., Bertram, R., & Pollatsek, A. (2004). Are long compound words identified serially via their constituents? Evidence from an eye-movement-contingent display change study. *Memory & Cognition*, *32*, 523–532.

Hyönä, J., & Olson, R. K. (1995). Eye movement patterns among dyslexic and normal readers: Effects of word length and word frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1430–1440.

Hyönä, J., & Pollatsek, A. (1998). Reading Finnish compound words: Eye fixations are affected by component morphemes. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 1612–1627.

Inhoff, A. W. (1984). Two stages of word processing during eye fixations in the reading of prose. *Journal of Verbal Learning and Verbal Behavior*, *23*, 612–624.

Inhoff, A. W. (1989). Lexical access during eye fixations in reading: Are word access codes used to integrate lexical information across interword fixations? *Journal of Memory and Language*, *28*, 444–461.

Inhoff, A. W., & Liu, W. (1997). The perceptual span during the reading of Chinese text. In: H. -C. Chen (Ed.), *The cognitive processing of Chinese and related Asian languages*. Hong Kong: The Chinese University Press.

Inhoff, A. W., & Liu, W. (1998). The perceptual span and oculomotor activity during the reading of Chinese sentences. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 20–34.

Inhoff, A. W., Pollatsek, A., Posner, M. I., & Rayner, K. (1987). Covert attention and eye movements during reading. *Quarterly Journal of Experimental Psychology*, *41A*, 63–89.

Inhoff, A. W., Radach, R., & Heller, D. (2000). Complex compounds in German: Interword spaces facilitate segmentation but hinder assignment of meaning. *Journal of Memory and Language*, *42*, 23–50.

Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics*, *40*, 431–439.

Ishida, T., & Ikeda, M. (1989). Temporal properties of information extraction in reading studied by a text–mask replacement technique. *Journal of the Optical Society A: Optics and Image Science*, *6*, 1624–1632.

Jared, D., Levy, B. A., & Rayner, K. (1999). The role of phonology in the activation of word meanings during reading: Evidence from proofreading and eye movements. *Journal of Experimental Psychology: General*, *128*, 219–264.

Johnson, R. E. (2006). The flexibility of letter coding: Nonadjacent letter transposition effects in the parafovea. In: R. P. G. von Gompel, M. H. Fischer, W. S. Murray, R. C. Hill (Eds), *Eye movements: A window on mind and brain*, Oxford: Elsevier.

Johnson, R. E., Perea, M., & Rayner, K. (2006). Transposed-letter effects in reading: Evidence from eye movements and parafoveal preview. *Journal of Experimental Psychology: Human Perception and Performance* (in press).

Juhasz, B. J. (2005). Age-of-acquisition effects in word and picture identification. *Psychological Bulletin*, *131*, 684–712.

Juhasz, B. J. (2006). The Influence of semantic transparency on eye Movements during English compound word recognition. In: R. P. G. von Gompel, M. H. Fischer, W. S. Murray, R. C. Hill (Eds), *Eye movements: A window on mind and brain,* Oxford: Elservier.

Juhasz, B. J., Inhoff, A. W., & Rayner, K. (2005). The role of interword spaces in the processing of English compound words. *Language and Cognitive Processes*, *20*, 291–316.

Juhasz, B. J., Liversedge, S. P., White, S. J., & Rayner, K. (2006). Binocular coordination of the eyes during reading: Word frequency and case alternation affect fixation duration but not binocular disparity. *Quarterly Journal of Experimental Psychology* (in press).

Juhasz, B. J., & Rayner, K. (2003). Investigating the effects of a set of intercorrelated variables on eye fixation durations in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 1312–1318.

Juhasz, B. J., & Rayner, K. (2005). Age of acquisition effects on eye movements in reading. *Visual Cognition*, *13*, 846–863.

Juhasz, B. J., Starr, M. S., Inhoff, A. W., & Placke, L. (2003). The effects of morphology on the processing of compound words: Evidence from naming, lexical decisions and eye fixations. *British Journal of Psychology*, *94*, 223–244.

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

Kambe, G. (2004). Parafoveal processing of prefixed words during eye fixations in reading: Evidence against morphological influences on parafoveal preprocessing. *Perception & Psychophysics*, *66*, 279–292.

Kambe, G., Rayner, K., & Duffy, S. A. (2001). Global context effects on processing lexically ambiguous words. *Memory & Cognition*, *29*, 363–372.

Kennedy, A. (2000). Parafoveal processing in word recognition. *Quarterly Journal of Experimental Psychology*, *53A*, 429–456.

Kennedy, A, & Pynte J. (2005). Parafoveal-on-foveal effects in normal reading. *Vision Research*, *45*, 153–168.

Kennison, S. M. (2001). Limitations on the use of verb information during sentence comprehension. *Psychonomic Bulletin & Review*, *8*, 132–138.

Kennison, S. M. (2002). Comprehending noun phrase arguments and adjuncts. *Journal of Psycholinguistic Research*, *31*, 65–81.

Kennison, S. M., & Clifton, C. (1995). Determinants of parafoveal preview benefit in high and low working memory capacity readers: Implications for eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 68–81.

Kliegl, R., & Engbert, R. (2003). SWIFT explorations. In: J. Hyönä, R. Radach, & H. Deubel (Eds) *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 391–412). Amsterdam: Elsevier.

Kliegl, R., & Engbert, R. (2005). Fixation durations before word skipping in reading. *Psychonomic Bulletin & Review*, *12*, 132–138.

Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology*, *16*, 262–284.

Legge, G. E., Hooven, T. A., Klitz, T. S., Mansfield, J. S., & Tjan, B. S. (2002). Mr. Chips 2002: New insights from an ideal-observer model of reading. *Vision Research*, *42*, 2219–2234.

Legge, G. E., Klitz, T. S., & Tjan, B. S. (1997). Mr. Chips: An ideal-observer model of reading. *Psychological Review*, *104*, 524–553.

Lima, S. D. (1987). Morphological analysis in reading. *Journal of Memory and Language*, *26*, 84–99.

Lima, S. D., & Inhoff, A. W. (1985). Lexical access during eye fixations in reading: Effects of word-iniital letter sequences. *Journal of Experimental Psychology: Human Perception and Performance*, *11*, 272–285.

Liversedge, S. P., & Findlay, J. M. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Science*, *4*, 6–14.

Liversedge, S. P., Pickering, M. J., Branigan, H. P., & van Gompel, R. P. G. (1998). Processing arguments and adjuncts in isolation and context: The case of by-phrase ambiguities in passives. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 461–475.

Liversedge, S. P., Pickering, M., Clayes, E. L., & Brannigan, H. P. (2003). Thematic processing of adjuncts: Evidence from an eye-tracking experiment. *Psychonomic Bulletin & Review*, *10*, 667–675.

Liversedge, S. P., Rayner, K., White, S. J., Vergilino-Perez, D., Findlay, J. M., & Kentridge, R. W. (2004). Eye movements when reading disappearing text: is there a gap effect in reading? *Vision Research*, *44*, 1013–1024.

Liversedge, S., White, S.J., Findlay, J.M., & Rayner, K. (2006). Binocular coordination of eye movements during reading. *Vision Research*, *46*, 2363–2374.

McConkie, G. W., Kerr, P. W., Reddix, M. D., & Zola, D. (1988). Eye movement control during reading: I. The location of initial eye fixations in words. *Vision Research*, *28*, 1107–1118.

McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics*, *17*, 578–586.

McConkie, G. W., & Zola, D. (1979). Is visual information integrated across successive fixations in reading? *Perception & Psychophysics*, *25*, 221–224.

McDonald, S. A., Carpenter, R. H. S., & Shillcock, R. C. (2005). An anatomically-constrained, stochastic model of eye movement control in reading. *Psychological Review*, *112*, 814–840.

McDonald, S. A., & Shillcock, R. C. (2003a). Eye movements reveal the on-line computation of lexical probabilies during reading. *Psychological Science*, *14*, 648–652.

McDonald, S. A., & Shillcock, R. C. (2003b). Low-level predictive inference in reading: The influence of transitional probabilities on eye movements. *Vision Research*, *43*, 1735–1751.

Meseguer, E., Carreiras, M., & Clifton, C. (2002). Overt reanalysis strategies and eye movements during the reading of mild garden path sentences. *Memory & Cognition*, *30*, 551–561.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing words: Evidence of a dependence upon retrieval operations. *Journal of Experimental Psychology*, *90*, 227–234.

Miellet, S., & Sparrow, L. (2004). Phonolgocial codes are assembled before word fixation: Evidence from boundary paradigm in sentence reading. *Brain and Language*, *90*, 299–310.

Miller, B., Juhasz, B. J., & Rayner, K. (2006). The orthographic uniqueness point and eye movements during reading. *British Journal of Psychology*, *97*, 191–216.

Morris, R. K. (1994). Lexical and message-level sentence context effects on fixation times in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 92–103.

Morris, R. K., & Folk. J. R. (1998). Focus as a contextual priming mechanism in reading. *Memory & Cognition*, *26*, 1313–1322.

Morris, R. K., Rayner, K., & Pollatsek, A. (1990). Eye movement guidance in reading: The role of parafoveal letter and space information. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 268–281.

Morrison, R. E. (1984). Manipulation of stimulus onset delay in reading: Evidence for parallel programming of saccades. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 667–682.

Morrison, R. E., & Rayner, K. (1981). Saccade size in reading depends upon character spaces and not visual angle. *Perception & Psychophysics*, *30*, 395–396.

Murray, W. S., & Rowan, M. (1998). Early, mandatory, pragmatic processing. *Journal of Psycholinguistic Research*, *27*, 1–22.

Myers, J. L., Cook, A. E., Kambe, G., Mason, R. A., & O'Brien, E. J. (2000). Semantic and episodic effects on bridging inferences. *Discourse Processes*, *29*, 179–200.

Ni, W., Crain, S., & Shankweiler, D. (1996). Sidestepping garden paths: Accessing 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*, 515–539.

Niswander-Klement, E., & Pollatsek, A. (2006). The effects of root-frequency, word frequency, and length on the processing of prefixed English words during reading. *Memory & Cognition,* 34, 685–702.

Nuthmann, A., Engbert, R., & Kliegl, R. (2005). Mislocated fixations during reading and the inverted optimal viewing position effect. *Vision Research*, *45*, 2201–2217.

O'Brien, E. J., Raney, G. E., Albrecht, J., & Rayner, K. (1997). Processes involved in the resolution of explicit anaphors. *Discourse Processes*, *23*, 1–24.

O'Brien, E. J., Shank, D. M., Myers, J. L., & Rayner, K. (1988). Elaborative inferences during reading: Do they occur on-line? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 410–420.

O'Regan, J. K. (1979). Eye guidance in reading: Evidence for the linguistic control hypothesis. *Perception & Psychophysics*, *25*, 501–509.

O'Regan, J. K. (1980). The control of saccade size and fixation duration in reading: The limits of linguistic control. *Perception & Psychophysics*, *28*, 112–117.

Paterson, K. B., Liversedge, S. P., & Underwood, G. (1999). The influence of focus operators on syntactic processing of short relative clause sentences. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *52A*, 717–737.

Pickering, M. J., & Frisson, S. (2001). Processing ambiguous verbs: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 556–573.

Pickering, M. J., & Traxler, M. J. (2001). Strategies for processing unbounded dependencies: Lexical information and verb-argument assignment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 1401–1410.

Pickering, M. J., & Traxler, M. J. (2003). Evidence against the use of subcategorisation frequencies in the processing of unbounded dependencies. *Language and Cognitive Processes*, *18*, 469–503.

Pollatsek, A., & Hyönä, J. (2005). The role of semantic transparency in the processing of Finnish compound words. *Language and Cognitive Processes*, *20*, 261–290.

Pollatsek, A., Hyönä, J., & Bertram, R. (2000). The role of morphological constituents in reading Finnish compound words. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 820–833.

Pollatsek, A., Lesch, M., Morris, R. K., & Rayner, K. (1992). Phonological codes are used in integrating information across saccades in word identification and reading. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 148–162.

Pollatsek, A., Raney, G. E., LaGasse, L., & Rayner, K. (1993). The use of information below fixation in reading and in visual search. *Canadian Journal of Experimental Psychology*, *47*, 179–200.

Pollatsek, A., & Rayner, K. (1982). Eye movement control in reading: The role of word boundaries. *Journal of Experimental Psychology: Human Perception and Performance*, *8*, 817–833.

Pollatsek, A., Rayner, K., & Balota, D. A. (1986). Inferences about eye movement control from the perceptual span in reading. *Perception & Psychophysics*, *40*, 123–130.

Pollatsek, A., Reichle, E., & Rayner, K. (2003). Modeling eye movements in reading. In: J. Hyönä, R. Radach, H. Deubel (Eds). *The mind's eyes: Cognitive and applied aspects of eye movement research* (pp. 361–390). Amsterdam: Elsevier.

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the E–Z Reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology*, *52*, 1–56.

Pynte, J., Kennedy, A., & Ducrot, S. (2004). The influence of parafoveal typographical errors on eye movements in reading. *European Journal of Cognitive Psychology*, *16*, 178–202.

Raney, G. E., & Rayner, K. (1995). Word frequency effects and eye movements during two readings of a text. *Canadian Journal of Experimental Psychology*, *49*, 151–172.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology*, *7*, 65–81.

Rayner, K. (1977). Visual attention in reading: Eye movements reflect cognitive processes. *Memory & Cognition*, *4*, 443–448.

Rayner, K. (1978). Eye movements in reading and information processing. *Psychological Bulletin*, *85*, 618–660.

Rayner, K. (1979). Eye guidance in reading: Fixation locations within words. *Perception*, *8*, 21–30.

Rayner, K. (1995). Eye movements and cognitive processes in reading, visual search, and scene perception. In: J. M. Findlay, R. Walker, & R. W. Kentridge (Eds), *Eye movement research: Mechanisms, processes and applications* (pp. 3–12). Amsterdam: North-Holland.

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

Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. (2004). The effects of frequency and predictability on eye fixations in reading: Implications for the E–Z Reader model. *Journal of Experimental Psychology: Human Perception and Performance*, *30*, 720–732.

Rayner, K., Balota, D. A., & Pollatsek, A. (1986). Against parafoveal semantic processing during eye fixations in reading. *Canadian Journal of Psychology*, *40*, 473–483.

Rayner, K., & Bertera, J.H. (1979). Reading without a fovea. *Science*, *206*, 468–469.

Rayner, K., Binder, K. S., Ashby, J., & Pollatsek, A. (2001). Eye movement control in reading: Word predictability has little influence on initial landing positions in words. *Vision Research*, *41*, 943–954.

Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behavior*, *22*, 358–374.

Rayner, K., Cook, A. E., Juhasz, B. J., & Frazier, L. (2006). Immediate disambiguation of lexically ambiguous words during reading: Evidence from eye movements, British Journal of Psychology (in press).

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition*, *14*, 191–201.

Rayner, K., & Duffy, S. A. (1988). On-line comprehension processes and eye movements in reading. In: M. Daneman, G. E. MacKinnon, & T. G. Waller (Eds), *Reading research: Advances in theory and practice* (pp. 13–66). New York: Academic Press.

Rayner, K., & Fischer, M. (1996). Mindless reading revisited: Eye movements during reading and scanning are different. *Perception & Psychophysics*, *58*, 734–747.

Rayner, K., Fischer, M. H., & Pollatsek, A. (1998). Unspaced text interferes with both word identification and eye movement control. *Vision Research*, *38*, 1129–1144.

Rayner, K., & Frazier, L. (1987). Parsing temporarily ambiguous complements. *The Quarterly Journal of Experimental Psychology*, *39A*, 657–673.

Rayner, K., & Frazier, L. (1989). Selection mechanisms in reading lexically ambiguous words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 779–790.

Rayner, K., Inhoff, A. W., Morrison, R. E., Slowiaczek, M. L., & Bertera, J. H. (1981). Masking of foveal and parafoveal vision during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance*, *7*, 167–179.

Rayner, K., & Juhasz, B. J. (2004). Eye movements in reading: Old questions and new directions. *European Journal of Cognitive Psychology*, *16*, 340–352.

Rayner, K., Juhasz, B. J., Ashby, J., & Clifton, C. (2003). Inhibition of saccade return in reading. *Vision Research*, *43*, 1027–1034.

Rayner, K., Kambe, G., & Duffy, S. A. (2000). Clause wrap-up effects on eye movements during reading. *Quarterly Journal of Experimental Psychology*, *53A*, 1061–1080.

Rayner, K., Li, X., Juhasz, B. J., & Yan, G. (2005). The effect of word predictability on the eye movements of Chinese readers. *Psychonomic Bulletin & Review*, *12*, 1089–1093.

Rayner, K., & Liversedge, S. P. (2004). Visual and linguistic processing during eye fixations in reading. In: J. M. Henderson, & F. Ferreira (Eds), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 59–104). Hove, UK: Psychology Press.

Rayner, K., Liversedge, S. P., & White, S. J. (2006). Eye movements when reading disappearing text: The importance of the word to the right of fixation. *Vision Research*, 46, 310–323.

Rayner, K., Liversedge, S. P., White, S. J., & Vergilino-Perez, D. (2003). Reading disappearing text: Cognitive control of eye movements. *Psychological Science*, *14*, 385–389.

Rayner, K., & McConkie, G. W. (1976). What guides a reader's eye movements. *Vision Research*, *16*, 829–837.

Rayner, K., McConkie, G. W., & Zola, D. (1980). Integrating information across eye movements. *Cognitive Psychology*, *12*, 206–226.

Rayner, K., Pacht, J. M., & Duffy, S. A. (1994). Effects of prior encounter and global discourse bias on the processing of lexically ambiguous words: Evidence from eye fixations. *Journal of Memory and Language*, *33*, 527–544.

Rayner, K., & Pollatsek, A. (1981). Eye movement control during reading: Evidence for direct control. *Quarterly Journal of Experimental Psychology, 33A*, 351–373.

Rayner, K., Pollatsek, A., & Binder, K. S. (1998). Phonological codes and eye movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 476–497.

Rayner, K., Raney, G. E., & Pollatsek, A. (1995) Eye movements and discourse processes. In: R. F. Lorch, & E. J. O'Brien (Eds), *Sources of coherence in reading* (pp. 9–36). Hillsdale, NJ: Lawsence Erlbaum Associates, Inc.

Rayner, K., & Sereno, S. C. (1994a). Eye movements in reading: Psycholinguistic studies. In: M. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 57–82). New York: Academic Press.

Rayner, K., & Sereno, S. C. (1994b). Regressive eye movements and sentence parsing: On the use of regression-contingent analyses. *Memory & Cognition*, *22*, 281–285.

Rayner, K., & Sereno, S. C. (1994c). Regression-contingent analyses: A reply to Altmann. *Memory & Cognition*, *22*, 291–292.

Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R., & Clifton, C. (1989). Eye movements and on-line comprehension processes. *Language and Cognitive Processes*, *4*, SI21–SI49

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.

Rayner, K., Reichle, E. D., Stroud, M. J., Williams, C. C., & Pollatsek, A. (2006). The effect of word frequency, word predictability, and font difficulty on the eye movements of young and elderly readers. *Psychology and Aging*, (in press).

Rayner, K., Warren, T., Juhasz, B. J., & Liversedge, S. P. (2004). The effect of plausibility on eye movements in reading. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *30*, 1290–1301.

Rayner, K., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review*, *3*, 504–509.

Rayner, K., Well, A. D., Pollatsek, A., & Bertera, J. H. (1982). The availability of useful information to the right of fixation in reading. *Perception & Psychophysics*, *31*, 537–550.

Rayner, K., White, S. J., Kambe, G., Miller, B., & Liversedge, S.P. (2003). On the processing of meaning from parafoveal vision during eye fixations in reading. In: J. Hyona, R. Radach, & H. Deubel (Eds), *The mind's eye: Cognitive and applied aspects of eye movements* (pp. 213–234). Oxford: Elsevier.

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.

Reichle, E. D., Pollatsek, A., & Rayner, K. (2006a). Modeling the effects of lexical ambiguity on eye movements during reading. In: R. P. G. von Gompel, M. H. Fischer, W. S. Murray, & R. C. Hill, (Eds), *Eye movements: A window on mind and brain*. Oxford: Elsevier.

Reichle, E. D., Pollatsek, A., & Rayner, K. (2006b). E–Z Reader: A cognitive-control, serial-attention model of eye-movement behavior during reading. *Cognitive Systems Research*, 7, 34–55.

Reichle, E. D., Rayner, K., & Pollatsek, A. (1999). Eye movement control in reading: Accounting for initial fixation locations and refixations within the E–Z reader model. *Vision Research*, *39*, 4403–4411.

Reichle, E. D., Rayner, K., and Pollatsek, A. (2003). The E–Z Reader model of eye movement control in reading: Comparison to other models. *Brain and Behavioral Sciences*, *26*, 445–476.

Reilly, R., & Radach, R. (2003). Foundations of an interactive activation model of eye movement control in reading. In: J. Hyönä, R. Radach, & H. Deubel (Eds), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 429–455). Oxford: Elsevier.

Reingold, E., & Rayner, K. (2006). Examining the word identification stages identified by the E–Z Reader model. *Psychological Science* (in press).

Schmauder, A. R., & Egan, M. C. (1998). The influence of semantic fit on on-line sentence processing. *Memory & Cognition*, *26*, 1304–1312.

Schustack, M. W., Ehrlich, S. F., & Rayner, K. (1987). Local and global sources of contextual facilitation in reading. *Journal of Memory and Language*, *26*, 322–340.

Sereno, S. C. (1995). Resolution of lexical ambiguity: Evidence from an eye movement priming paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 582–595.

Sereno, S. C., O'Donnell, P., & Rayner, K. (2006). Eye movements and lexical ambiguity resolution: Investigating the subordinate bias effect. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 335–350.

Sereno, S. C., Pacht, J. M., & Rayner, K. (1992). The effect of meaning frequency on processing lexically ambiguous words: Evidence from eye fixations. *Psychological Science*, *3*, 296–300.

Sereno, S. C., & Rayner, K. (2000). Spelling-sound regularity effects on eye fixations in reading. *Perception & Psychophysics*, *62*, 402–409.

Sparrow, L., & Miellet, S. (2002). Activation of phonological codes during reading: Evidence from errors detection and eye movements. *Brain and Language*, *81*, 509–516.

Speer, S. R., & Clifton, C. Jr. (1998). Plausibility and argument structure in sentence comprehension. *Memory & Cognition*, *26*, 965–978.

Starr, M. S., & Rayner, K. (2001). Eye movements during reading: Some current controversies. *Trends in Cognitive Science*, *5*, 156–163.

Staub, A., & Rayner, K. (2006). Eye movements and comprehension processes. In: G. Gaskell (Ed.), *Oxford Handbook of Psycholinguistics.* Oxford: Oxford University Press.

Sturt, P. (2003). The time course of the application of binding constraints in reference resolution. *Journal of Memory and Language*, *48*, 542–562.

Tinker, M. A. (1939). Reliability and validity of eye-movement measures of reading. *Journal of Experimental Psychology*, *19*, 732–746.

Traxler, M. J., McElree, B., Williams, R. S., & Pickering, M. J. (2005). Context effects in coercion: Evidence from eye movements. *Journal of Memory and Language*, *53*, 1–25.

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

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

Traxler, M. J., Pickering, M. J., & McElree, B. (2002). Coercion in sentence processing: Evidence from eye movements and self-paced reading. *Journal of Memory and Language*, *47*, 530–547.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic disambiguation. *Journal of Memory and Language*, *33*, 285–318.

van Gompel, R. P. G., & Majid, A. (2004). Antecedent frequency effects during the processing of pronouns. *Cognition*, *90*, 255–264.

van Gompel, R. P. G., Pickering, M. J., Pearson, J., & Liversedge, S.P. (2005). Evidence against competition during syntactic ambiguity resolution. *Journal of Memory and Language*, *52*, 284–307.

van Gompel, R. P. G., Pickering, M. J., & Traxler, M. J. (2001). Reanalysis in sentence processing: Evidence against current constraint-based and two stage models. *Journal of Memory and Language*, *45*, 225–258.

Vitu, F. (1991). The influence of parafoveal processing and linguistic context on the optimal landing position effect. *Perception & Psychophysics*, *50*, 58–75.

Vitu, F., McConkie, G. W., Kerr, P., & O'Regan, J. K. (2001). Fixation location effects on fixation durations during reading: an inverted optimal viewing position effect. *Vision Research*, *41*, 3513–3533.

White, S. J., Rayner, K., & Liversedge, S.P. (2005). The influence of parafoveal word length and contextual constraint on fixation durations and word skipping in reading. *Psychonomic Bulletin & Review*, *12*, 466–471.

Wiley, J., & Rayner, K. (2000). Effects of titles on the processing of text and lexically ambiguous words: Evidence from eye movements. *Memory & Cognition*, *28*, 1011–1021.

Williams, R. S., & Morris, R. K. (2004). Eye movements, word familiarity, and vocabulary acquisition. *European Journal of Cognitive Psychology*, *16*, 312–339.

Yan, G., Tian, H., Bai, X., & Rayner, K. (2006). The effect of word and character frequency on the eye movements of Chinese readers. *British Journal of Psychology*, *97*, 259–268.

Yang, S., & McConkie, G. W. (2001). Eye movements during reading: A theory of saccade initiation times. *Vision Research*, *41*, 3567–3585.

Yang, S., & McConkie, G. W. (2004). Saccade generation during reading: Are words necessary? *European Journal of Cognitive Psychology*, *16*, 226–261.