Rayner, K., & Clifton, C., Jr. (2002). Language processing. In D. Medin (Volume Editor) Stevens Handbook of Experimental Psychology, Third Edition: Volume 2, Memory and Cognitive Processes (pp 261-316). New York: John Wiley and Sons, Inc.

Language Comprehension

Keith Rayner and Charles Clifton, Jr.

Department of Psychology

University of Massachusetts

Amherst, MA 01003

Correspondence to:
Keith Rayner
Department of Psychology
University of Massachusetts
Amherst, MA 01003
413-545-2175; rayner@psych.umass.edu

In this chapter, we discuss a number of important phenomenon with respect to language comprehension. We must acknowledge that to completely review all aspects of language processing would be a task that could hardly be accomplished in an entire book, let alone a single chapter. So, undoubtedly we will leave out some peoples' favorite topic in language processing (for more detailed coverage see Garrod & Pickering, 1999; Gernsbacher, 1994). We will discuss both reading and listening with an eye towards reviewing what is known about each domain. The chapter consists of four main sections: (1) tasks and paradigms, (2) comprehending words, (3) comprehending sentences, and (4) comprehending text. In the first section, we briefly review the tasks and paradigms that have been used to study language processing. In each of the remaining sections, we discuss (a) the nature of the task, (b) the core phenomena, and (c) representations and models of the specific process.

Tasks and Paradigms

The goal of experimental psychologists who study language processing is to discover how the complex processes of the mind operate when language is understood. To do so, a number of tasks and paradigms have been developed to observe, record, interpret, and predict the activity of the mind. In this section, we will discuss a number of such tasks and paradigms that have been utilized to study language comprehension processes. Specifically, we will delineate how each has been used to examine how people extract meaning from both written and spoken language. Since most of these techniques have been shown to have both strengths and weaknesses, we will also discuss some of the limitations inherent in various tasks (see Haberlandt, 1994 for a more complete discussion of various tasks used to study language processing).

Reaction time measures.

Reaction time measures are arguably the most common procedure for tapping into comprehension processes, and psycholinguists generally use such measures to examine the relative time-course of a process. Reaction time (RT) is defined as the interval between the presentation of a stimulus and the onset of the subject's subsequent response. This interval is typically measured with a high degree of precision (e.g., in milliseconds), and response types vary from simply naming the stimulus to making a more complex decision, such as deciding whether two words are related in meaning to one another.

Naming, lexical decision, and categorization. In the naming task, subjects are asked to articulate a word or a pronounceable nonword and reaction times are measured from the presentation of the stimulus to the onset of the named response. By contrast, in the lexical decision task, subjects must decide whether a letter string is a word (e.g., desk) or a nonword (e.g., dosk), with reaction times measured from the presentation of the letter string to the onset of the word/nonword response. A third task is categorization, in which subjects must judge whether or not a given word belongs to some predetermined category (Is it a living thing?). For the most part, in our discussions below, we will focus on results from naming and lexical decision since these tasks have been used more frequently than categorization to study language processing.

In the past, the most popular usage of naming and lexical decision has been to determine the time-course of visual word identification. For example, when factors such as word length and syntactic class are controlled, naming and lexical decision times for high frequency (more common) words are shorter than those for low frequency (less common) words. However, one problem with such tasks is that overall response time is not simply a measure of word identification, since both naming and lexical decision times also include the time it takes a subject to formulate and initiate the appropriate response for the task (i.e., an articulation or a manual button-press). Furthermore, it is not clear whether such responses even require the subject to identify the stimulus. The naming task, for example, simply permits the subject to pronounce a string of letters based upon grapheme-to-phoneme conversion rules without necessarily requiring that word meaning be accessed (e.g., most people can formulate a pronunciation for blicket, although no corresponding meaning exists in the dictionary). Similarly, in the lexical decision task, subjects may be able to judge whether a letter string is a word by simply basing their decision on the familiarity of the letter string rather than on the actual identification of the word. This does not necessarily mean that these tasks are insensitive to semantic properties of words — quite the contrary, naming and lexical decision tasks have been shown to exhibit effects of word frequency and familiarity, which would be unlikely unless some aspect of word meaning was accessed. Despite these limitations, response times in these tasks may be used to classify the upper limits of the time course for word recognition. However, many researchers have used naming and lexical decision tasks in conjunction with other tasks to determine whether the patterns of reaction times converge (see Taft, 1991).

Priming and masking. Two methodologies have emerged which are often used in conjunction with naming and lexical decision. The priming paradigm (Meyer & Schvaneveldt, 1971) involves the presentation of a sequence of two words: a prime then a target. Subjects are asked to make a decision regarding the target word, and how quickly they are able to respond is measured. An early finding that emerged from priming studies is that when the prime is semantically related to the target (e.g., dog followed by cat), subjects respond more quickly than when the prime is not semantically related to the target (e.g., pen followed by cat). This indicates that the relationship between prime and target words influences processing time on the target.

A second paradigm, masking, also examines word identification time by limiting the exposure of a stimulus. For example, the word dog is presented for 60 ms (and then disappears) and is replaced by a pattern mask consisting of either a series of x's, random letters, or letter-like shapes. Although reaction time is often measured in masking studies, the more common procedure is to measure accuracy. As with many reaction time measures, the masking paradigm has been useful in allowing researchers to examine the time course of lexical processing. Studies utilizing masking techniques also suggest that subjects may extract information from words which are presented for such a brief duration (e.g., less than 30 ms) that they are not aware of the prime words identity (Balota, 1983; Marcel, 1983).

More recently, a paradigm which combines priming and masking procedures, masked priming (Forster & Davis, 1984), has been used to shed light on the early stages of word comprehension. In this paradigm, subjects look at a fixation target and a mask is presented followed by the brief presentation of a prime word which is then followed by a target word (presented for about 200 ms) which is in turn followed by another mask. Although subjects are generally unable to identify the prime word, it still has an effect on their report of the target word.

Dual tasks and phoneme monitoring. The dual task paradigm is often used to study attentional processes, and it follows two basic assumptions: (1) that subjects have a limited processing capacity and (2) that different cognitive activities may make use of different processing resources. For example, subjects may be asked to read sentences while listening for a tone. If response times are slower when performing two tasks simultaneously as compared to performing a single task (e.g., simply reading sentences), this would be evidence that both tasks are drawing upon the same cognitive resources. Further, the rate of slowdown may also indicate the degree of resource utilization.

One example of the dual task paradigm is phoneme monitoring. Most often, phoneme monitoring is used to study speech comprehension, and it involves listening to auditorily presented sentences while monitoring for a particular phoneme (e.g., to detect the /b/ sound while listening to the sentence The emperor went to the royal baths.). Thus the two tasks are to comprehend the sentence and to press a button when the target phoneme is detected. The idea is that if contextual or lexical processing prior to the target word (e.g., baths in the example) is difficult, it should take subjects longer to detect a target phoneme. For example, subjects are slower to detect a phoneme when the target word is preceded by an ambiguous word. Researchers utilizing this task are often interested in determining the basic units of speech perception or in studying the processing complexity of sentence contexts, lexical ambiguity, and attentional issues. However, the data emerging from phoneme monitoring tasks are often affected by a number of extraneous variables such as the frequency of targets across sentence stimuli, the discriminability of the phoneme, target word length, and the frequency of the target word in which the to-be-detected phoneme is located.

Speed accuracy tradeoff. In addition to the dual task paradigm, another technique puts subjects under various types of speed constraint. For example, one variation of the technique involves training subjects to respond immediately upon the presentation of a signal that occurs at various times after the end of a sentence. Accuracy of a decision made about a sentence increases as the response deadline increases. The parameters of the function with which accuracy increases as the response deadline increases can reveal information about processing activities (McElree, 1993). The major concern with this technique is that it may induce strategies that are specific to the demands of the task.

Processing time and other measures to assess comprehension.

The reaction time measures discussed above are most commonly used when the unit of interest is a single stimulus (e.g., a word). Researchers interested in examining readers' comprehension of larger units, such as sentences or sentence phrases, use one of a variety of processing time methodologies. By manipulating characteristics of the text, researchers can infer selected attributes of comprehension processes. In these tasks, subjects may be asked to read a paragraph or sentence while elapsed reading times are recorded (so that paragraph reading time or sentence reading time is measured). Similarly, subjects may be given a limited amount of time to read a portion of text while error rates are recorded.

Self-paced reading and listening. Sometimes, if an experimenter is interested in how long it takes a subject to read a particular segment of text, measuring overall reading times for sentences or paragraphs may be too imprecise. In the self-paced reading task, the experimenter controls the amount of text that the subject can see at any one time, and the size of the segment (e.g., a word or a phrase) available to the subject is generally a function of the topic under investigation. When the subject has finished reading one segment, s/he pushes a button and the next segment of text is presented. When only one word at a time is presented, this procedure yields a processing time measure for each word in the text (Just, Carpenter, & Woolley, 1982). A variation of this task is called the "stops making sense" task, in which subjects advance word-by-word through a sentence as long as it makes sense; when the sentence no longer makes sense, or becomes ungrammatical, subjects push a different button (Boland, Tanenhaus, & Garnsey, 1990).

One problem with the self-paced reading task is that it does not mimic natural reading. Reading times in the self-paced reading paradigm are slower (about half as fast) than those in more natural reading tasks since subjects must press a button to read subsequent segments of text. Since it takes longer to manually press a button than it does to move the eyes, words stay on the screen for about 400 ms in this task, as compared to average eye fixation times of approximately 250 ms in natural reading. Given that reading in the self-paced paradigm is slower in general, one possibility is that subjects may develop different comprehension strategies.

More recently, self-paced listening paradigms (Ferreira, Henderson, Anes, Weeks, & McFarlane, 1996) have been developed to study speech perception. Just as in the reading situation, the listener pushes a button to get the next word or segment of discourse. Similar concerns regarding strategic effects also apply to this paradigm.

RSVP. Natural silent reading involves moving the eyes to successive segments of text—hence the reader controls how quickly text is read. By contrast, in the rapid serial visual presentation (RSVP) task (Potter, Kroll, & Harris, 1980), the experimenter controls the rate at which text is presented. In this paradigm, the subject sits in front of a computer screen while new words are presented one at a time for various durations (e.g., 50 to 400 ms). Studies utilizing this technique have found that readers can comprehend short passages of text which are presented at rates of up to 1,200 words per minute, with a new word being presented every 50 ms. Interestingly, when each word is presented for 250 ms, reading comprehension in the RSVP task is often better than in natural reading. However, this paradigm has critical limitations. Although comprehension performance is high for short passages of text, as the amount of text increases, comprehension begins to suffer (Masson, 1983). This is partially because RSVP reading prevents readers from looking back at “misunderstood” portions of text (during normal reading, readers make move their eyes back to previously read text on approximately 10% of all eye fixations). Moreover, RSVP reading is also mentally taxing for subjects, as it requires their constant attention to text.

Phoneme restoration. The self-paced reading and RSVP tasks described are generally utilized to measure higher-order cognitive comprehension processes in reading. In contrast, the phoneme restoration effect has most commonly been used to measure lower-order, perceptual processing in listening. The phoneme restoration effect is an auditory illusion that arises when part of an utterance is replaced by an extraneous sound such as a cough or white noise. In such instances, listeners often perceptually fill in (restore) the missing phoneme and report that they heard the complete utterance (Warren, 1970). Early studies using this method found that psychoacoustic factors related to the nature of the replacement sound (e.g., amplitude and quality) affected the probability of detecting the missing phoneme (Warren & Obusek, 1971). Subsequent studies have used the restoration effect to examine the extent to which lexical and higher-level representations can influence speech perception (Samuel, 1981, 1996). Samuel (1996) notes that, while effects emerging from the phoneme restoration paradigm are real, the paradigm is sensitive to small changes in methodology, e.g., differences in the syllabic length of the carrier word, the phonological class of the replaced segment, and the quality of the replacing sound.

Eye movements

With the continuing development of technological innovations, some researchers have begun to replace or supplant reaction time and processing time paradigms with eye movement measures. In a typical eye-tracking experiment, subjects read sentences presented on a computer monitor while their eye movements are recorded. Researchers then look at patterns of readers' eye movements noting, for example, how long readers' eyes remain fixated on words or phrases within sentences, how far their eyes move from fixation to fixation, or how frequently their eyes regress back to re-read text.

Eye movements have been utilized to study a variety of language comprehension processes, and data gleaned from eye-tracking studies have been found to reflect moment-to-moment cognitive processes. One early finding was that where readers look and how long they look there is directly related to the ease or difficulty of cognitive processing (see Rayner, 1978, 1998). For example, when extraneous factors are controlled, fixation times are longer for lower frequency words, which are less likely to be encountered during reading, as compared to higher frequency words, which are more likely to be encountered. Eye movements have also been used to examine the effects of lexical ambiguity, morphological complexity, discourse processing, semantic relatedness, phonological processing, syntactic disambiguation, and the perceptual span (see Rayner, 1998; Rayner & Sereno, 1994, for reviews).

Eye-movement contingent display changes. A number of methods have emerged within the eye-tracking paradigm including the development of the eye-movement contingent display change paradigm (see Figure 1). In this paradigm, text displayed on a computer screen is manipulated as a function of where the eyes are fixated. As readers' eyes move across a line of text, letters or words may be modified in foveal, parafoveal, or peripheral locations, thus allowing the experimenter to control the nature and amount of information available to the reader. One variation of the eye-movement contingent paradigm is the moving window paradigm (McConkie & Rayner, 1975; Rayner & Bertera, 1979). In this paradigm, as readers move their eyes across the text, upon each fixation, text is exposed within an experimenter defined "window" while all text outside of the window is altered in some way (e.g., all of letters might be replaced by X's). Wherever the reader looks, the text within the window is available. The logic of the paradigm is that when the window is as large as the region from which information can normally be obtained, reading will proceed as smoothly as when there is no window (normal text). Using this technique, the size of the perceptual span in reading has been determined.

Insert Figure 1 about here

Another variation is the boundary paradigm (Rayner, 1975), in which characteristics of a target word in a particular location within a sentence may be manipulated. For example, in the sentence <u>John composed a new tune for the children</u>, when readers' eyes move past the space between <u>new</u> and <u>tune</u>, the target word <u>tune</u> would change to <u>song</u>. In this manner, researchers can examine the types of information (e.g., orthographic, phonological, semantic) that readers obtained from the target word prior to fixating upon it. Indeed, readers do process a target word more quickly (preview benefit) when they have had a preview of that word.

A final variation is the fast-priming paradigm (Sereno & Rayner, 1992), in which a prime word is briefly presented for a very short duration (i.e., less than 50 ms) and is immediately replaced by a target word. Primes may be related in meaning to target words (<u>tune</u>-<u>song</u>), but they may also be phonologically related (e.g., bat-cat) or orthographically related (e.g., bench-beach). This paradigm has been used to examine the time course of word processing. An advantage of using eye-movement measures over reaction time and processing time measures is that they allow researchers to study comprehension processes in a more natural setting. As mentioned previously, one disadvantage of reaction time and processing time measures is that they may result in the formulation of task-specific strategies or may simply slow the reading process. In the eye-movement paradigm, readers are free to read text as they would during normal reading. Moreover, eye movement measures are flexible, allowing researchers to examine both fine-grain and coarse-grain language comprehension processes.

Eye movements and listening. Eye movement recording techniques have also been utilized in the context of speech understanding. It has been demonstrated that when subjects listen to a narrative while a scene is presented in front of them which depicts objects in the narrative, their eyes tend to move to those objects that are mentioned in the narrative. This technique, often called the head-mounted eyetracking technique, allows researchers to make inferences about on-line speech comprehension (Tanenhaus & Spivey-Knowlton, 1996).

Physiological measures

In the past 20 years, a number of physiological measures have been developed to study cognitive processes. These measures range from simply recording heart rate to recording more complex physiological activity, such as measuring changes in brain activity. It is hoped that by using such measures, researchers will be able to accomplish two major goals: (1) to locate language comprehension regions (or pathways) in the brain and (2) to more closely examine the time course of cognitive activity within the brain. Although there are many physiological measures, in this section, we will focus only on those measures which involve examining activity in the brain (see Gazzaniga, 2000 for a more complete review of physiological measures).

ERP. Among the most common physiological measures used today is the event-related potential or ERP (Kutas & Van Petten, 1994) which involves measuring electrical events in the brain using electrodes placed on the scalp. By averaging electrical potentials over a number of trials, researchers hope to time-lock brain activity to a particular sensory event (e.g., the presentation of a word stimulus). The voltages associated with brain activity vary in both polarity and magnitude over time, resulting in a series of electrical "peaks and valleys". For example, when subjects are presented with a semantic incongruity, a relatively large negative potential (i.e., a valley) occurs about 400 ms after the presentation of the stimulus (this is termed a N400 wave).

One advantage of using ERP's over other methodologies is that they allow experimenters to more directly examine the time course of language comprehension processes within the brain itself. On the other hand, there is no guarantee that the ERP activity being measured is the direct result of a particular cognitive process, as opposed to being the result of later (e.g., memory) processing. While ERPs have very good temporal resolution, much of the research has focused on late occurring waves. One problem here for reading is that since various effects occur within 250 ms (as evidenced by eye movement data), the events reflected in the late occurring ERP signal take place after the relevant processing activities have occurred. So, if a certain effect shows up during an eye fixation (less than 250-300 ms), examining the same effect in the N400 doesn't seem to provide direct information about the time course of the effect (Raney & Rayner, 1993; Sereno & Rayner, 2000a). Thus, it may be worthwhile to examine earlier occurring ERP waves than has typically been the case (Sereno, Rayner, & Posner, 1999).

PET. Positron-emission tomography (PET) scans (Petersen, Fox, Posner, Mintum, & Raichle, 1989) are based on a different framework than ERP's. This method involves the ingestion of a small amount of radioactive material which may be traced and used to measure blood flow in the brain; cognitive activity is indexed by changes in blood flow to active parts of the brain. Studies using PET scans have found that many different parts of the brain are involved in language comprehension (including parts of the left temporal, parietal, and frontal cortex).

This complexity is perhaps the greatest disadvantage to the PET methodology. It is not surprising that language comprehension involves the coordination of a number of brain systems, but the metabolic activity measured by PET scans may also reflect additional processing not directly related to language. For example, researchers have found increased metabolic activity in brain systems which are not specific to language processing per se. Specifically, studies examining reading processes have found increased metabolic activity in the anterior cingulate cortex, which is normally associated with sustained attentional processing, as well as in the contralateral cerebellum, which is thought to be involved in the rapid shifting of attention.

MRI/*f*MRI. Magnetic resonance imaging, or MRI, and its newest counterpart, functional magnetic resonance imaging, *f*MRI (Buckner, 1998), are based on framework similar to that of a PET scan—namely, that sensory, motor, and cognitive tasks produce a localized increase in neural activity which gives rise to subsequent increases in blood flow. In very general terms, MRI is based upon how cells that are relatively rich or poor in oxygen respond to a magnetic field; fMRI reflects in changes in blood flow while a subject is engaged in a cognitive task. Researchers utilizing MRI technology to examine language processes are typically interested in localizing language comprehension functions in the brain. For example, in a baseline condition an experimenter may present subjects with a word and simply require the subject to look at the word. In another condition, subjects may be asked to decide whether the word represents a living thing. Differences in neural activity between the two conditions can then be used to determine the region in the brain used in processing aspects of word meaning.

One advantage of the MRI/*f*MRI paradigm is that it is relatively non-invasive and represents little health risk to subjects (as opposed to PET scans which involve the ingestion of potentially harmful radioactive materials). In addition, they permit the experimenter to collect hundreds (or even thousands) of images from a single subject, with highly accurate spatial resolution. MRI technology is also becoming increasingly available to psycholinguists, as many hospitals have MRI facilities.

The MRI/*f*MRI paradigm also suffers from several disadvantages. The most significant limitation is that temporal resolution is relatively poor (e.g., although it may only take about 250 ms to recognize a word, an *f*MRI can only acquire data in about 2-3 seconds), thus disallowing any clear examination of the time course of language processing. However, some scientists have also begun to combine the temporal resolution of ERPs with the spatial resolution of *f*MRI's. In addition, as mentioned earlier in reference to PET scans, a great deal of activity in the brain occurs which is only indirectly related to language functions, resulting in some degree of difficulty in localizing areas in the brain specific to language comprehension.

We have presented the physiological based methods for the sake of completeness. However, the bulk of the research discussed in this chapter comes from methodologies and paradigms other than the brain imaging methods (e.g., PET and fMRI). The reason for this is quite simple - research on language comprehension using brain imaging is only in its infancy and at this point there are very few imaging studies that really elucidate our understanding of language processing. To be sure, such techniques have revealed a great deal about which brain regions are active during various types of language processing, just not that much about processing per se.. However, we expect that in the near future many such studies will appear.

Word Recognition

The nature of the task

Clearly, recognizing the individual words in texts and discourse represents the first stage for understanding language. Some would undoubtedly argue that grapheme (letter) or phoneme (the smallest sound unit) recognition, as well as morpheme (the smallest meaningful unit) recognition necessarily must precede word recognition and there have been lively research efforts in pursuit of understanding the recognition of these three units. However, given space limitations, we will focus first on the word before moving to sentence and discourse comprehension.

Considerable effort has been devoted to understanding how words are recognized during reading and listening. If language comprehension consisted only of recognizing individual words, the task for researchers interesting how language is understood would be considerably easier than it is. But, words do not occur in isolation and exactly how individual words are integrated into a discourse representation is an interesting question. Much of the research on visual word recognition has focused on how readers access the meaning of a word. In traditional models of word recognition, meanings of words are represented in the reader's lexicon (or mental dictionary where information about a word, such as its meaning is stored). In these models (Fig

2), there are two pathways, one from graphemic units to meaning directly, and one from graphemic units to phonological units, and then to meaning (the phonological mediation pathway). In this Dual Route Model (Coltheart, 1978; Coltheart, Curtis, Atkins, & Haller, 1993, Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001), the direct pathway can be used for words that become highly familiar and must be used to read so called "exception words" (e.g., café) for which an indirect phonological route would fail. And the phonological route must be used to read pseudowords (e.g. nufe) for which there is no lexical representation to access. These issues of mediation and one-or-two routes are central points of contrast between traditional representational models and more recent alternative theoretical models.

Insert Figure 2 about here

These alternative models share the idea that words are not represented in a mental lexicon, but rather emerge from processing activity. Connectionist models (Fig 3) are nonrepresentational in that they assume that words emerge from patterns of parallel and distributed activation (Harm & Seidenberg,1999; Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg & McClelland, 1989). Resonance models are also nonrepresentational, claiming that word identification results from the stabilization of dynamic patterns that are continuously modified by interactions among inputs and various dynamic states resulting from prior experience (Stone & Van Orden, 1994;Van Orden & Goldinger,1994). An interesting feature of this model is that patterns of graphic-phonological activation stabilize more rapidly than do patterns of graphic-semantic activation. In effect, a word form is identified primarily through the convergence of orthography and phonology. Meaning is slower to exert an influence on the identification process.

Insert Figure 3 about here

These contrasting approaches to cognitive architectures have been central in word recognition research in recent years. Although there may not be enough data for deciding between representational and nonrepresentational models (e.g. Besner, Twilley, McCann, & Seergobin, 1990; Seidenberg & McClelland, 1989; Coltheart, et al, 1990), both classes of models can account for many of the same results. Most models of word identification have not generally been sensitive to the linguistic structure of words. In both representational and nonrepresentational models, a letter string has usually been treated like any other stimulus. And its spoken language counterpart, to which the graphic stimulus must be connected, has been treated like any other associated response. One recent model, however, specifies a functional internal linguistic structure during the word identification process. In this Two-cycles model (Berent & Perfetti, 1995), vowels and consonant phonemes are assembled separately from a graphic input. Experiments using masking and a variation of the fast priming paradigm have demonstrated that, in English, consonants are assembled more quickly than vowels (Berent & Perfetti, 1995; Lee, Rayner, & Pollatsek, 2001). This separation of vowels and consonants can arise either from fundamental phonological considerations or from the fact that, in English at least, the grapheme-phoneme mapping is more reliable for consonants than for vowels.

In this section, we will discuss visual and auditory word recognition, and focus on some issues that have received the attention in the literature. Due to space limitations, we will not discuss some important topics, such as neighborhood effects (Andrews, 1997) and cross linguistic studies (Frost, 1998; Lukatela & Turvey, 1998). We will begin by discussing some relations between visual and auditory word recognition. Then, we discuss two topics which have received considerable attention in the visual word identification literature. First, we discuss whether visual word identification is a letter by letter or parallel process. Second, we discuss the extent to which sound codes are involved in visual word identification. Following our discussion of these topics, we turn to discussions of frequency effects, context effects, and ambiguity effects.

Visual and auditory word recognition

We start with the basic assumption that at some level of the processing system, input from the eyes and from the ears converge on a central processing system. However, listening and reading also differ in seemingly fundamental ways. In listening, the information speech provides about what words are being uttered is spread over time. Identification of a heard word is a serial process, in which information arriving at the beginning of the word is used earlier, and differently, than information arriving later in the word (Cutler, Dahan, & Van Donselaar, 1997; Cutler & Clifton, 2000). In written language, information is spread over space, not over time. Reading, though based upon auditory language, permits a substantial amount of parallel processing of the visual word form. The eyes land on a word and apparently acquire information from various parts of a word in parallel (Rayner & Pollatsek, 1989). Some serial processing is induced by multiple fixations made on a word, but the visual word recognition process seems to be primarily a parallel process (see the following section for further discussion).

This apparently-fundamental difference between reading and listening has led to the development of quite different models of visual and auditory word recognition. Models of visual word recognition concentrate on questions like whether or not distinct representations (e.g., visual and phonological) are computed in the process of word recognition and on how different sources of information (visual, lexical, contextual, etc.) are integrated. In contrast, most current models of auditory word recognition have the flavor of Marslen-Wilson's Cohort model (Marslen-Wilson & Welsh, 1978). A listener activates an initial set, or cohort, of words that are consistent with the beginning segments of an auditory word. This cohort is reduced by information from later segments and by competition between the members of the cohort, hopefully resulting in the recognition of a single word. There is substantial debate about whether top-down feedback from the lexicon guides perception of individual speech sounds (the TRACE model of McClelland & Elman, 1986 claims it does; the MERGE model of Norris, McQueen, & Cutler, 2000, claims it does not). What is clear is that multiple candidate words are activated during the process of recognition: Words that are embedded within a spoken word or overlap with it are momentarily activated.

This multiple activation brings up the question of how the beginning and end of a word is identified. This is a non-problem in reading, at least in languages that are written like English: Words are separated by spaces, and readers use these spaces (Rayner, 1998). But in listening, there are no spaces. Words run into one another, more or less seamlessly. It appears that the form of language provides some cues to the starts of words. For instance, languages generally have phonotactic constraints on what sequences of segments can occur together in a word. Violating these constraints means that a word boundary has been crossed. Further, some languages provide useful prosodic cues to the beginnings of words. English content words, for example, tend to begin with a strong syllable, and listeners seem to treat the occurrence of a strong syllable as evidence that a new word may have begun (see Cutler et al., 1997). Beyond formal cues to the segmentation of words from the speech stream, it appears that success in recognizing a word is an important signal in dividing the speech stream into words (Cutler & Clifton, 2000). For instance, possible words like stay, steak, and take may be activated during the perception of the string first acre, but eventually success in exhaustively assigning the auditory signal to these two words will inhibit the competing overlapping words.

One problem that is common to both listening and reading is determining the nature of the lexical entry that is contacted by an auditory or a visual word. A major point of theoretical dispute is whether the representation that is stored in the lexicon consists of a stem together with information about its possible affixes (so, for example, a lexical entry would be read together with information about the possible affixes re-, -er, -ing-, etc.) vs. a full form (e.g. reading) together with links to related full forms. Marslen-Wilson, Tyler, Waksler, & Older (1994) propose the former position; McQueen and Cutler (1998) advocate the latter. Although the nature of the lexical entry has not been completely determined, it is clear that in English recognizing a spoken or written word makes available morphological information such as part of speech and number marking that can be used in later stages of comprehension.

Serial versus parallel processing. An interesting and important issue with respect to how words are recognized in reading deals with whether they are processed as wholes (in parallel) or letter-by-letter (serially). More than 100 years ago, Cattell (1886) addressed this question by briefly exposing words and letters and asking people to report what they saw. In fact, subjects were better able to report the words than the letters. However, there were several well-known flaws in the experiment having to do with guessing and memory artifacts.

Insert Figure 2 about

Cattell's experiment lay dormant until Reicher (1969) and Wheeler (1970) replicated the experiment but with important controls to eliminate the possible artifacts. The characteristics of their paradigm are shown in Figure 1. Basically, a word, single letter, or non-word letter string was presented very briefly (about 25-40 ms) and followed immediately by a masking pattern which would interfere with any extended processing of the stimulus after its offset. Next, two letter choices were presented: the correct letter and a letter that was not presented. Notice, however, that the alternative letter in the word condition was always such that it made a word. Thus, although the letter d was actually presented, the alternative response letter k also made a word. The basic finding in these experiments (and many subsequent experiments) was that a letter is better identified when it is embedded in a word than when it is presented in isolation or in a non-word. This result, the word superiority effect, suggests that Cattell's phenomenon is real: letters in words are identified more accurately than letters in isolation. The phenomenon forces two conclusions. First, the serial letter-by-letter view of word recognition can't be correct: it should take much longer to process four letters than one and given that the mask disrupts processing, if this view were correct subjects should be more accurate with single letters than words. Second, all of the letters in the word are processed since accuracy at identifying the correct letter was independent of the letter position tested.

In general, the word superiority effect has been taken as evidence that all of the letters in a word are processed in parallel. The effect ultimately led to the development of powerful computer simulation models designed to account for the results. These models, called the Interactive Activation model (McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982) and the Verification model (Paap, Newsome, McDonald, & Schvaneveldt, 1982) were the forerunners of even more elegant and powerful connectionist models (e.g., Seidenberg &

McClelland, 1989) mentioned earlier.

Sound Coding in Word Identification. So far, we have discussed visual word identification as if it was a purely visual process. However, given that alphabets are designed to code for the sounds of the words, it seems plausible that the process of identifying words is not purely visual and also involves accessing the sounds that the letters represent and possibly assembling them into the sound of a word. Moreover, if we think about accessing the sound of a word, it becomes less clear what "word identification" means. Is it accessing a sequence of abstract letters, accessing the sound of the word, accessing the meaning of the word, or some combination of all three? In addition, what is the casual relationship between accessing the three types of codes? One possibility is that readers merely access the visual code -- more-or-less like getting to a dictionary entry -- and then "look up" the sound of the word and the meaning in the lexicon. Another relatively simple possibility is that, for alphabetic languages, readers first access the sound of the word and only then, access the meaning. That is, in this view, the written symbols merely serve to access the spoken form of the language and meaning is tied only to the spoken form. On the other hand, the relationship may be more complex. For example, the written form may start to activate both the sound codes and the meaning codes, and then the three types of codes send feedback to each other to arrive at a "solution" as to what the visual form, auditory form, and meaning of the word are. There are probably few topics in reading that have generated as much controversy as this: what the role of sound coding is in the reading process.

Given that naming words is quite rapid (around 500 ms) and given that a significant part of this time must be taken up in both programming the motor response and executing the motor act of speaking, it seems plausible that accessing the sound code could be rapid enough to be part of the process of getting to the meaning of a word. But even if accessing the sound code is accessed at least as rapidly as the meaning, it may not play any causal role. Certainly, there is no logical necessity for involving the sound codes, as the sequence of letters is sufficient to access the meaning (or meanings) of the word and in the Interactive Activation model (McClelland & Rumelhart, 1981) and the Verification (Paap et al., 1982) models, access to the lexicon (and hence word meaning) is achieved via a direct look-up procedure which only involves the letters which constitute a word. However, before examining the role of sound coding in accessing the meanings of words, let's first look at how sound codes, themselves, are accessed.

The Access of Sound Codes. There are three general possibilities for how we access the pronunciation of a letter string. Many words in English have irregular pronunciations (e.g., <u>one</u>), such that their pronunciations cannot be derived from the spelling-to-sound rules as defined by the language. In these cases, it would seem that the only way to access the sound code would be via a direct access procedure, where the word's spelling is matched to a lexical entry within the lexicon. In the above example, the letters <u>o-n-e</u> would activate the visual word detector for one which would in turn activate the subsequent lexical entry. Once this entry is accessed, the appropriate pronunciation for the word (/wun/) could be activated. In contrast, other words have regular pronunciations (e.g., <u>won</u>). Such words' pronunciations could also be accessed via a direct route, but their sound codes could also be constructed through the utilization of spelling-to-sound correspondence rules or by analogy to other words in the language (Glushko, 1981). Finally, it is, of course, possible to pronounce nonwords like <u>mard</u>. Unless all possible pronounceable letter strings have lexical entries (which seems unlikely), nonwords' sound codes would have to be constructed.

Research on acquired dyslexics, who were previously able to read normally but suffered brain damage making reading even single words difficult, has revealed two constellations of symptoms that seem to argue for the existence of both a direct and a constructive route to a word's pronunciation (Coltheart, Patterson, & Marshall, 1980). In surface dyslexia, the patients can pronounce both real words and nonwords but they tend to "regularize" irregularly pronounced words (e.g., pronouncing island as /iz-land/). In contrast to surface dyslexics, deep and phonemic dyslexics can pronounce real words (whether they are regular or irregular), but they cannot pronounce nonwords. It was initially believed that surface dyslexics completely relied on their intact constructive route, whereas deep dyslexics completely relied on their direct route. However, it is now realized that these syndromes are somewhat more complex than they were first thought to be, and the descriptions of the syndromes are somewhat oversimplified. Nonetheless, they do seem to argue that the two processes (a direct look-up process and a constructive process) may be somewhat independent of each other.

Assuming that these two processes exist in skilled readers (who can pronounce both irregular words and nonwords correctly) how do they relate to each other? The simplest possibility is that they operate independently of each other in a "race". Whichever process finishes first would presumably "win" and determine the pronunciation. Thus, since the direct look-up process can't access the pronunciation of nonwords, the constructive process would determine the pronunciation for nonwords. What would happen for words? Presumably, the speed of the direct look-up process would be sensitive to how frequent the word was in the language, with low frequency words taking longer to access than high frequency words. However, the constructive process, which is not dependent on lexical knowledge, should be largely independent of the word's frequency. Thus, for common (i.e. frequent) words, the pronunciation of both regular and irregular words should be determined by the direct look-up process and should take more-or-less the same time. For less frequent words, however, both the direct and constructive process would be operating because the direct access process would be slower. Thus, for irregular words, there would be conflict between the pronunciations generated by the two processes. Irregular words should be pronounced more slowly (if the conflict is resolved successfully) or with errors if the word is "regularized".

The data from a number of studies are consistent with such a "race" model. A very reliable finding (see Baron & Strawson, 1976; Perfetti & Hogaboam, 1975) is that regular words are pronounced (named) more quickly than irregular words. However, the difference in naming times between regular and irregular words is a function of word frequency: for high frequency words there is little or no difference, but there are large differences for low frequency words. However, the process of naming is likely to be more complex than a simple race, as people usually make few errors in naming, even for low frequency irregular words. Thus, somehow, the two routes "cooperate" in some way to produce the correct pronunciation, but when the two routes conflict in their output, there is slowing of the naming time.

There are two final points worth noting. First, few words are totally "irregular". That is, even for quite irregular words like <u>one</u> and <u>island</u>, the constructive route would produce a pronunciation that had some overlap with the actual pronunciation. Second, the fact that low frequency words take more processing time is not restricted to naming or lexical decision tasks. Sereno and Rayner (2000b) found that evidence for this frequency by regularity interaction in that readers' eye fixations were longer on irregular low frequency words than on irregular high frequency words.

Sound Codes and the Access of Word Meaning. In the previous section we discussed how readers access a word's sound codes. However, a much more important question is how a word's meaning is accessed. As indicated earlier, this has been a highly contentious issue with respected researchers stating quite differing positions (with some claiming that readers do not form articulatory representations of printed words and others claiming that the heart of reading is the decoding of written symbols into speech). Although a great deal has been learned about this topic, the controversy represented by this dichotomy of views continues, and researchers' opinions on this question still differ greatly.

Some of the first attempts to resolve this issue utilized the lexical decision task. One question was whether there was a difference between regularly spelled words and irregularly spelled words in such a task, under the tacit assumption that this task was reflecting the speed of accessing the meaning of words (Bauer & Stanovich, 1980; Coltheart, 1978). Unfortunately these data tended to be highly variable in that some studies found a regularity effect while others did not. Meyer, Schvaneveldt, and Ruddy (1974) utilized a somewhat different paradigm, and found that the time for subjects to determine whether touch was a word was slower when it was preceded by a word such as couch (which presumably primed the incorrect pronunciation) as compared to when it was preceded by an unrelated word. However, there is now considerable concern that the lexical decision is fundamentally flawed as a measure of lexical access that is related to accessing a word's meaning. The most influential of these arguments was that it induces artificial checking strategies before making a response (Balota & Chumbley, 1984, 1985; Chumbley & Balota, 1984).

A task that presumably does get at the issue of how a word's meaning is accessed is the categorization task. As noted earlier, in this task, people are given a category label (e.g., tree) and then are given a target word (e.g., beech, beach, or bench) and have to decide whether it represents a member of the preceding category (Van Orden, 1987; Van Orden, Johnston, & Hale, 1988; Van Orden, Pennington, & Stone, 1990). The key finding was that subjects had a hard time rejecting homophones of true category exemplars (e.g., beach). Not only were they slow in rejecting these items, they typically made 10-20% more errors on these items than control items that were visually similar (e.g., bench). In fact, these errors persisted even under conditions when people were urged to be cautious and go slowly. Moreover, the effect is not restricted to word homophones. A similar, though somewhat smaller, effect was reported with pseudohomophones (like brane). Moreover, in a similar semantic relatedness judgment task (i.e., decide whether the two words on the screen are semantically related), subjects are slower and make more errors on false homophone pairs such as pillow-bead. (Bead is a "false homophone" of pillow because bead could be a homophone of bed analogously to head rhyming with bed.) These findings (Lesch & Pollatsek, 1998) with pseudohomophones and false homophones both make it unlikely that these results are merely due to people just not knowing the spelling of the target words and argue that phonology plays a significant role in accessing a word's meaning.

Sound codes in reading. The work that we described above has dealt largely with the extent to which sound codes are involved in identifying words in isolation. However, there has also been considerable research on the use of sound codes in reading sentences and extended discourse. Some research (Slowiaczek & Clifton, 1980) suggested that sound codes are useful in that they aid comprehension processes since information about prosodic structure may be available from them. The more critical question, however, is whether or not sound codes are activated early or late in processing during reading. To address this issue, Pollatsek, Lesch, Morris, and Rayner (1992) utilized the boundary paradigm discussed earlier to examine whether phonological codes are active before words are even fixated (and hence very early in processing). In their study, the preview word was either identical to the target word (beach or shoot), a homophone of it (beech or chute), or an orthographic control word (bench or shout). Thus, readers were presented with a different word in the target word location before they fixated it and during the saccade to that region a display change occurred (so that beech changed to beach or chute changed to shoot), although they were not aware of the change. The key finding was that fixation time on the target word was shorter when the preview was a homophone than when it was just orthographically similar. This indicates that, in reading text, sound codes are extracted from words even before they are fixated, which is quite early in the encoding process. Research by Rayner, Sereno, Lesch, and Pollatsek (1995; see also Lee, Binder, Kim, Pollatsek, & Rayner, 1998; Lee, Rayner, & Pollatsek, 1999) using the fast priming technique has also obtained evidence that sound codes are activated very early during reading. Other standard eye movement experiments have also found effects consistent with this claim (Folk & Morris, 1995; Folk, 1999).

Some other paradigms, however, have come up with less positive evidence for the importance of sound coding in word identification during reading. Daneman and Reingold (1993; Daneman, Reingold, & Davidson, 1995) used a manipulation in a reading study similar to the preview study with three conditions: correct homophone, incorrect homophone, and spelling control (e.g., Alone at his teller's cage, idle and -- with the sentence continuing with either bored, board, or beard). However, in their studies, a "wrong" word (either the wrong homophone or the spelling control) remained in the text throughout the trial. People read short passage containing these "errors" and the key question was whether the wrong homophones would be less disruptive than the spelling controls because they "sounded right". In two studies, Daneman and colleagues found there was disruption in the reading process (measured by examining the time spent looking at the target word before moving to another word) for both types of wrong words, though no significant difference between the wrong homophones and the spelling control, though they did find more disruption for the spelling control when later measures of processing (which included regressive fixations on the target word were examined). This finding is consistent with a view where sound coding plays only a back-up role in word identification. However, in a similar type of study, Rayner, Pollatsek, and Binder (1998) found greater disruption for the spelling control than for the wrong homophone on "immediate" measures of processing. Even in the Rayner et al. study, the homophone effects are relatively small (far more so than in the Van Orden paradigm mentioned above). It appears that sentence and paragraph context may interact with word processing to make errors (be they phonological or orthographical) less damaging to the reading process (see Jared, Levy, & Rayner, 1999 on the issue of interactions with reading skill).

Recognizing Words in Sentences

Frequency effects. In both reading and listening, it is clearly the case that the frequency of the word matters. Numerous experiments have demonstrated that high frequency words are processed faster and more efficiently than low frequency words. Before discussing some typical results, let’s back up a bit and ask the question “How is frequency determined”? Most typically, it is done by relying on corpus data in which the occurrence of each word in the corpus is tabulated and summed. The most frequently used corpus data are the Francis and Kučera (1982) and CELEX (Baayen, Piepenbrock, & van Rijn, 1993) counts. These data bases typically provide information not only about how frequently each word form is used but also provide information about usage (e.g., is the word used as a noun or verb). The frequency of a word is usually measured by taking some corpus of text that is assumed to be representative and actually counting the number of times that a particularly word occurs. To give some feel for frequency counts, words such as irk, jade, cove, vane, and prod have counts of 2 to 5 per million (in the Francis and Kučera count), while words like cat, coat, greet, and square all have frequencies greater than 30 per million. The most frequent words are typically short function words like the, and, of, and for. Thus, there is a confounding in natural language such that the most frequent words are typically shorter than other words and they also tend to be function words (as opposed to content words). However, in typical experiments dealing with word frequency effects, these very high frequent function words would not be used. Rather, words would be matched on extraneous variables (such as word length, number of syllables, part of speech, etc) so that the only difference between the high and low frequency words is indeed their frequency of occurrence.

The difference in lexical decision time between a high frequency word like coat and a low frequency word like cove is about 100-150 ms, while the difference in naming times is more like 30-60 ms (Balota & Chumbley, 1984, 1985; Chumbley & Balota, 1984). Clearly, both of these differences can't be estimates of how much longer it takes to identify a low frequency word than a high frequency word. When people are asked to read sentences containing high or low frequency target words (with the same neutral context preceding the target word, so that every word in the sentence is the same except for the target word), the difference in fixation time between them is closer to 30-60 ms (Inhoff & Rayner, 1986; Rayner & Duffy, 1986; Rayner, Sereno, & Raney, 1996). Thus, 30-60 ms seems like a better guess as to the effect of frequency in reading; the longer time associated with the lexical decision task has generally been assumed to be because there is a time consuming decision stage in that task. One interesting additional observation from the eye movement studies is that not only is the fixation time longer on the target word, it is also longer on the next word. Thus, the processing of the low frequency word apparently “spills over” onto the processing of the next word in the sentence. Schilling, Rayner, and Chumbley (1998) recently specifically compared readers’ performance in the three tasks and found that naming and fixation times yielded similar frequency effects for the target words whereas the effect on lexical decision time was much larger. They also found that there was a correlation between the frequency effect and average response time in the three tasks. In general, Schilling et al.’s data suggest that both naming and lexical decision tasks yield data about word recognition processes that are consistent with effects found in eye fixations in silent reading.

Frequency effects have also been observed in categorization tasks (Lewellen, Goldinger, Pisoni, & Greene, 1993; Monsell, Doyle, & Haggard, 1989) in which people have to decide whether or not a word belongs to a certain category and in the early components of the ERP response (Sereno, Rayner, & Posner, 1999). In listening experiments, frequency effects have also been observed (Ferreira et al., 1996; Foss & Blank, 1980). In reading, a particularly interesting effect reported by Henderson and Ferreira (1990) is that when readers fixate on a low frequency word, they are able to obtain less information from the word to the right of fixation in parafoveal vision than when they are fixated on a high frequency word (see also Kennison & Clifton, 1995).

Although frequency effects are rather ubiquitous in language processing research, there is some controversy about whether effects that have been attributed to word frequency are actually due to an age of acquisition factor that is correlated with frequency (Ellis & Lambon Ralph, 2000; Morrison & Ellis, 1995). According to this view, what is really critical is the age at which a word was learned rather than frequency of occurrence per se. On the other hand, a recent study (Lewis, Gerhand, & Ellis, 2001) demonstrated that cumulative frequency effects can account for age-of-acquisition effects. While we agree that this issue is important, we suspect that more data need to be obtained before final determinations can be made concerning the relative weight to give to each factor. Likewise, it is also sometimes argued (Gernsbacher, 1984; Lewellen et al., 1993) that familiarity ratings in which people rate how familiar a given word is might be more important than corpus frequency measures. Again, we do not disagree that familiarity per se is an important factor. However, what strikes us as most important in the present context is that words that are more frequent, more familiar, and acquired earlier are easier to process when reading or listening.

What exactly do frequency effects reflect? There is some controversy as to whether frequency effects reflect lexical access or integration processes. According to an integration account, high frequency words would be easier to integrate into the discourse representation that a reader is building than a low frequency word. In general, we suspect that most researchers believe that frequency effects are due to lexical access processes. One prominent model of eye movements during reading, the E-Z Reader model (Reichle, Pollatsek, Fisher, & Rayner, 1998), certainly associates frequency effects with access processes. The fact that there are spillover effects (Rayner & Duffy, 1986) from reading a low frequency word suggests that frequency does affect integration processes, but the frequency effect is primarily due to access processes.

Contextual constraint effects. Another variable known to influence the processing of a word is the degree of contextual constraint (or predictability) for a given word. The first experiments (Tulving & Gold, 1963; Tulving, Mandler, & Baumal, 1964; Morton, 1964) attempting to demonstrate context effects in a reading-like situation involved having subjects read a sentence fragment like The skiers were buried alive by the sudden.... The subjects were then shown the target word avalanche very briefly. They were able to identify the target word at significantly briefer exposures when the context predicted it than when it was preceded by neutral, inappropriate, or no context. These results were assumed to demonstrate that context affects the identification of words during reading. However, many researchers have questioned whether such situations have any bearing on normal reading. The brief presentation virtually guarantees that the visual information from the target word is not fully registered and hence degraded. The identification of a word in this situation is thus likely to be the result of a slow conscious problem solving process rather than normal perceptual identification.

Accordingly, the procedure was subsequently modified so that the target word appeared after the sentence frame was presented until the participant made a response to it (Stanovich & West, 1983). Most of these experiments required the participant to either name the target word (Stanovich & West, 1979; Becker, 1985) or make a lexical decision regarding the target word (Schuberth & Eimas, 1977; Fischler & Bloom, 1979). While this type of procedure alters the natural reading process, the timing relations aren't too different from normal reading if the delay between the sentence context and the target word is relatively brief. In most of these experiments, it has been shown that a highly constraining context facilitates naming or lexical decision latency relative to a neutral condition such as the frame The next word in the sentence will be. We should note that there has been some controversy over the appropriate baseline to use in these experiments, but that is beyond the scope of this chapter.

While experiments in which a sentence fragment is presented followed by a target word approximate fluent reading, there are some key differences. The first, of course, is that an extraneous response (naming or lexical decision) is called for, and the second is that a delay is introduced between the sentence fragment and the target word. Since the best way to study fluent reading is through the use of eye movements, we will focus on how they have been employed to study context effects. There are now numerous experiments that demonstrate that contextual constraint has strong influences on eye movements in reading (Balota, Pollatsek, & Rayner, 1985; Ehrlich & Rayner, 1981; Morris, 1994). First, words that are highly constrained by the context are skipped more frequently than words that are not highly constrained (the difference is usually about 10-15%). This has often been argued to mean that the predictable words are identified on the prior fixation. Second, when the target words are not skipped, highly constrained words are fixated for less time than unconstrained words. Third, when the word to the right of fixation is highly constrained by the preceding context, readers get more preview benefit from it than when it is not so constrained (Balota et al., 1985). In general, the pattern of results obtained in these experiments suggests that context can speed lexical access. The fact that more preview benefit is obtained when the next word is highly constrained is consistent with this idea as is the fact that highly constrained words are skipped more than unconstrained words.

The eye movement studies establish that context does affect the amount of time spent on a word, although the effects are relatively modest (fixation time is generally about 20-25 ms shorter on constrained words than unconstrained words). They also make a "guessing game" model of reading quite unlikely. In guessing game models (Goodman, 1970; Hochberg, 1970; Levin & Kaplan, 1970), conscious prediction is seen as a major factor in identifying words in text so that visual information does not have to be fully processed. These types of models have not been very popular for some time since there seem to be too many costs associated with conscious prediction. First, such prediction is likely to be wrong more than it is right since readers are not very good at predicting the next word, even with unlimited amounts of time (Gough, Alford, & Holley-Wilcox, 1981). Second, one would expect that such prediction would take processing resources away from the higher-order processes needed to put the words together to form syntactic and semantic structures. Third, if guessing what the next word is in reading were an important part of word identification, one would expect much bigger effects than those obtained in the eye movement experiments where the deck has been stacked to make guessing prevail. In short, the data suggest that while context does affect processing time for a word, it is at a stage of speeding processing rather than due to some type of guessing.

Contextual constraint effects have also been demonstrated in listening experiments. Prior to discussing these effects, we must note that some theories of speech perception (e.g., McClelland & Elman, 1986) are generally highly interactive and permit higher level processes to influence what you actually hear. According to them, the speech sounds that you perceive may actually be different from the speech sounds that actually reach our ears because cognitive and contextual factors influence our perception of the actual speech signal. Evidence for this comes from the phonemic-restoration effect (Samuel, 1981; Warren, 1970; Warren & Warren, 1970).

Warren (1970) presented listeners with a recording of the sentence The state governors met with their respective legi*latures convening in the capital city, but a 120 ms portion of the sentence was replaced with a coughing sound (indicated by the asterisk in the example). Only 1 of 20 listeners reported detecting a missing sound replaced by a cough and that person misreported its location. The other 19 heard the word legislature as they restored the missing phoneme that is predicted by the contextual information. In a subsequent study, Warren and Warren (1970) demonstrated that listeners are capable of using a great deal of information to determine the correct sound. They presented listeners with the sentence It was found that the *eel was on the fragment, and then they heard either axle, show, orange, or table. All listeners heard the same sentence except for the last word and the asterisk again indicates a missing segment in the sentence. Depending on the word which they actually heard at the end of the sentence, listeners reported hearing wheel, heel, peel, or meal. Again, the context influenced the listener's report of a sound, although in this case the relevant context occurred so far after the missing phoneme that the restoration effect presumably reflected a report bias (cf. Saumel, 1981).

Another example that listeners use context to help perceive speech comes from studies reported by Marslen-Wilson (1973, 1975; Marslen-Wilson & Welsh, 1978) which used a "fast shadowing" paradigm in which people had to rapidly shadow (repeat) speech. Some distortions were made in the actual speech, so that listeners actually heard the pseudoword cigaresh. They were more likely to restore the distortion to the proper pronunciation (cigarette) if the word was highly predictable from the preceding context (for example, He still wanted to smoke a cigaresh).

It may be that, under some common circumstances, contextual information plays a larger role in listening than in reading. The speech signal may commonly be degraded, masked by other sounds, casually or quickly spoken; in such cases, we may rely more heavily on higher order processes to make sense of the signal. In reading, on the other hand, a visually degraded stimulus is unusual, and we do have the luxury to go back and look at information that has already been processed if the word was not adequately recognized or encoded initially. We doubt, though, that there is a principled difference between reading and listening in the extent to which a discourse context affects the encoding of individual words.

Ambiguity effects. Words can be ambiguous on a number of dimensions. They can be phonologically ambiguous in that two words can be pronounced the same way (but spelled differently, as in break-brake) and mean different things. There are also sense ambiguities (as in newspaper, which can mean what you read every morning or an organization). And, words can be lexically ambiguous in that two meanings are spelled and pronounced the same (as in bank and straw). The issue of how lexically ambiguous words are processed has been center-stage in debates concerning the extent to which various components of language are modular or interactive. That is, does language processing consist of a number of different modules, each selectively sensitive to a limited range of information? Or, is language processing a highly interactive process in which low level modules (involved in word recognition) are influenced by higher order processing (like effects due to contextual effects)?

The earliest experiments on lexical ambiguity were done with people listening to sentences with ambiguous words (Foss & Jenkins, 1973; Swinney & Hakes, 1976). Processing difficulty was measured using the phoneme monitoring task. The results of these early studies were consistent with exhaustive access (both meanings of the ambiguous word are activated when it is encountered, so phoneme monitoring time was slower following the ambiguous word) (but cf. Mehler, Sequi, & Carey, 1978; Newman & Dell, 1978, for some criticisms of the task). More direct evidence came from experiments (Swinney, 1979; Tanenhaus, Leiman, & Seidenberg, 1979) that used a cross-modal priming task to support exhaustive or at least multiple access. In the cross-modal priming task, a subject hears a target sentence such as The man found several insects, spiders, and other bugs in the room. The word bugs is ambiguous in that it can mean either an insect or a listening device. In the task, at some point relative to when subjects hear bugs (either before they hear it, simultaneously with hearing it, or after hearing it) a word appears on a video monitor and they must respond to it (via a lexical decision). The lexical decision times for three different kinds of words are then examined: ant (related to the insect meaning), spy (related to the other meaning), or sew (a word unrelated to either meaning (the length and frequency of the target words was controlled). The key question is whether the lexical decision time to spy is any faster than to the unrelated word sew. In fact, when the target word appeared on the monitor at the offset of the ambiguous word, the response times for spy and ant were about the same, and faster than the time for sew. This result clearly suggests that both meanings of an ambiguous word are accessed, even when the context strongly disambiguates the meaning of the word. It was also found that if the target word was delayed for about 200 ms, the priming effect on spy disappeared, indicating that context quickly inhibited the "wrong" meaning of bugs.

Other experiments replicated the basic result (Burgess, Seidenberg, & Tanenhaus, 1988; Onifer & Swinney, 1981; Seidenberg, Tanenhaus, Leiman, & Bienknowski, 1982) and extended it by demonstrating that both meanings of an ambiguous word are accessed (using priming as the indicator of access), even when one of the meanings is highly dominant such as boxer. However, in these cases, the less dominant meaning (type of dog) was no longer active after about 200 ms, even when the prior context did not disambiguate the word. Thus, it appears that rapid inhibition of one of the meanings is produced not only by prior context but also by a more dominant meaning of a lexical item inhibiting a weaker one. Further, there is some reason to think that a context that is sufficiently strongly biased can promote selective as opposed to multiple access (Tabossi, 1988) and that measures more sensitive than naming time (specifically, ERPs that reflect semantic anomaly) can provide evidence of selective access (Van Petten & Kutas, 1987; see also Simpson, 1984, 1994, for additional discussion of the controversy over the details of how lexically ambiguous words are processed in listening).

What about during reading? Although there are again issues related to certain details of the findings (see Rayner, Binder, & Duffy, 1999; Vu & Kellas, 1999), there seems to be clear evidence that both contextual information and meaning dominance play critical roles. There are now a large number of eye movement studies (see Duffy, Morris, & Rayner, 1988; Rayner & Duffy, 1986; Rayner & Frazier, 1989; Rayner, Pacht, & Duffy, 1994; Sereno, Pacht, & Rayner, 1992 for examples) that have examined how lexically ambiguous words are processed during reading. The basic findings from this research suggest that both meaning dominance and contextual information influence the processing of such words. For ambiguous words (like straw) where there are two equally likely meanings, readers look longer at such words in neutral contexts than at a control word matched in length and word frequency. However, when the prior context disambiguates the meaning that should be instantiated, fixation times are no longer than on the control word. Thus, the contextual information helps the reader choose the appropriate meaning. For ambiguous words (like bank) where one meaning is much more dominant than the other, in neutral contexts readers look no longer at the ambiguous word than the control word. On the other hand, if the following parts of the sentence make it clear that the subordinate meaning should be instantiated, fixation times on the disambiguating information are quite long and regressions back to the target word are frequent (suggesting that the reader incorrectly selected the dominant meaning and now has to recompute the subordinate meaning). Conversely, when the disambiguating information precedes the biased ambiguous word and the subordinate meaning is instantiated, readers' gaze durations on the ambiguous word are lengthened. Apparently, the contextual information increased the level of activation for the subordinate meaning so that the two meanings are in competition (just as the two meanings of a balanced ambiguous word like straw are in competition in a neutral context). This general pattern of results has been interpreted in the context of a model called the Reordered Access model (Duffy et al., 1988) and the data have been simulated using a constraint-satisfaction framework (Duffy, Kambe, & Rayner, 2001).

Comprehending sentences

The nature of the task.

Identifying the words in a written or spoken sentence is only the initial step toward arriving at an understanding of its meaning and function. It is a necessary step; it makes the word's sense and its other linguistic properties available. But understanding a sentence is far more than simply combining the meanings of its words. A reader or listener must determine how words are related to one another in phrases, how these phrases are related to one another, how these phrases and their relations are to be interpreted semantically, and how the semantic object that corresponds to the sentence is related to the ongoing discourse.

The property of language that makes all this possible is its compositionality. A reader or listener can compose the meaning of a complex object, like a sentence, out of the meanings of its parts. In doing so, the reader/listener is guided by knowledge of the grammar of the language, the principles that govern how words can be combined into sentences. From this perspective, the core question in the psychology of language comprehension is, how does the reader/listener use his or her knowledge of the grammar of his/her language to recover the intended meaning of sentences?

Two aspects of language make the task of "parsing" a challenging one. The first is that sentences are frequently ambiguous, at least momentarily (cf. Church & Patil, 1982, for a corpus-based analysis). This fact is illustrated by the sentences in Table 1 (from Frazier & Clifton, 1996). All the sentence forms illustrated in this table have been used in research designed to see how readers resolve temporary ambiguity. The second aspect of language that makes parsing difficult is that the words and phrases that are related to one another are not necessarily adjacent in a sentence. The existence of "long distance dependencies" in language follows from the fact that natural language permits one phrase or sentence to be included inside another, recursively. A relative clause structure (The boy whom I used to know left) provides a simple example, in which the verb left is dependent on the noun phrase (NP) the boy. Another structure that results in a long-distance dependency in English is the *wh*-question, illustrated in Who did you used to know, where the question NP *who* is dependent on the direct object position following know. Were it not for temporary ambiguities and long-distance dependencies, the phrase structure of a sentence could in principle be recovered by a very simple parsing mechanism. It is only because of these properties of natural language that the task is an interesting and challenging one.

Insert Table 1 about here

In this section, we will review empirical evidence that puts constraints on explanations of how people succeed in the task of comprehending written and spoken sentences (see Clifton & Duffy, 2001). We will place this evidence in the context of various theoretical interpretations of parsing, but refer readers elsewhere for extensive discussions of parsing theory (e.g., Frazier, 1990, 1995a; Frazier & Clifton, 1996; Gibson, 1998; Jurafsky, 1996; MacDonald, Pearlmutter, & Seidenberg, 1994; Mitchell, 1994; Tanenhaus & Trueswell, 1995; Vosse & Kempen, 2000).

Priority of grammar. It is clear that readers and listeners use their knowledge of language structure (morphology, syntax, semantics, and phonology) to constrain their interpretations of sentences. Early in the history of psycholinguistics, there was some debate about whether grammatical knowledge plays a role in normal sentence comprehension (cf. Fodor, Bever, & Garrett, 1974; Riesbeck & Schank, 1978). However, since then, a wide range of phenomena have pointed to the use of detailed grammatical knowledge in interpreting sentences. Frazier (1983) presented some elementary pre-experimental arguments that readers and listeners <u>must</u> use their knowledge of grammar. For instance, Frazier discussed how English speakers form dramatically different interpretations of the sentences <u>The umpire helped the child to third base</u> and <u>The umpire helped the child on third base</u>, and can form two very different interpretations of <u>He showed her baby pictures</u> (<u>He showed her baby the pictures</u> vs. <u>He showed her the baby pictures</u>). These interpretations are triggered by what might appear to be very minor changes in the lexical makeup of the sentences except that the changes have major implications for grammatical structure. Scarcely any experiments on sentence comprehension in the past two decades have suggested otherwise. To take just one example, the fact that the very first fixation on was discovered is disrupted while reading After you drank the strange looking water was discovered to be polluted (Frazier & Rayner, 1982) implies that readers take detailed information about grammatical possibilities into account. Presumably, readers first give the strange looking water the grammatical analysis of direct object of drank, but realize that this analysis was in error as soon as they read was.

Grammatical knowledge thus takes priority in the sense that it must be honored in sentence comprehension, even if doing so results in an implausible or incoherent message. After all, as Garrett (1976) noted, the purpose of grammar is to allow us to say surprising things. But it is an open question whether grammatical knowledge takes priority in a temporal or logical sense over other sources of information that might guide sentence interpretation. Some theorists (e.g. Frazier, 1978, 1990) suggest that it does; other theorists (e.g. MacDonald et al, 1994) suggest otherwise (and still others, most notably Bever, Sanz & Townsend, 1998, argue that syntax is used only after the use of semantics).

Immediacy. The past three decades of study of sentence and text comprehension strongly suggest that readers and listeners generally arrive at a semantic interpretation of a sentence in an apparently-incremental and nearly-immediate fashion. They do not wait for the end of a clause or sentence (as Fodor et al., 1974, suggested), but instead their understanding of a sentence seems to keep up with words as they are heard or as the eyes land on them. Marslen-Wilson's (Marslen-Wilson, 1973; Marslen-Wilson, 1975) rapid auditory shadowing studies demonstrated this, as do Pickering and Traxler's (Pickering & Traxler, 1998; Pickering, Traxler, & Crocker, 2000; Traxler & Pickering, 1996a) more recent eye movement studies. As noted earlier, Marslen-Wilson found that some mispronunciations (cigaresh) were spontaneously corrected more often when they made sense in a context than when there was no supporting context, and Pickering and Traxler showed that the implausibility of a word (magazine, in As the woman sailed the magazine....) led to a nearly-immediate disruption of eye movements (cf. also Clifton, 1993; Speer & Clifton, 1998).

It may be that not all semantic interpretation is incremental and immediate. Frazier and Rayner (1987) found that syntactic category ambiguities (is train a verb or a noun in The desert trains....?) are not resolved immediately, and Frazier and Rayner (1990) found the same thing about the sense of a word (newspaper – the paper object or the institution). However, as a first approximation, readers and listeners do seem to build interpretations on an on-line word-by-word basis.

Depth-first theories of parsing: Structural simplicity and beyond.

The earliest work in the modern psycholinguistics of sentence comprehension demonstrated effects of structural similarity on sentence memory (Clifton & Odom, 1966; Mehler, 1963) and effects of structural complexity on sentence comprehension and verification (Gough, 1965; Stolz, 1967). For example, multiply self-embedded sentences, such as The first shot the tired soldier the mosquito bit fired missed, are demonstrably hard to comprehend and remember. However, this early work did not support the development of plausible theories of how structural factors played a role in the sentence comprehension process.

Interesting theories of how the sentence comprehension mechanism uses a reader/listener's knowledge of grammar in comprehending sentences were developed by linguists and computer scientists in the 1970's (Ford, Bresnan, & Kaplan, 1982; Frazier, 1978; Kimball, 1973; Marcus, 1980). These theories attempted to identify the factors that a reader or listener would use to create an initial structural analysis of a sentence that could then be semantically interpreted and integrated into a discourse (but cf. Forster, 1979, for an early "parallel" model that assumed all structural analyses were initially created)

These theories, particularly Frazier's "garden path" theory, led to a great deal of empirical research and eventually to the creation of interesting alternative theories. Frazier's basic claim is that the reader/listener uses his/her knowledge of the syntactic structure of language, encoded as phrase structure rules or templates, to attach each incoming word into a phrase structure tree that represents the phrasal segments of a sentence and the relations between them. All phrase structure templates that contain the incoming word's syntactic category as a terminal element are examined in parallel. The first template (if there is one) whose root matches an accessible node in the existing phrase structure tree is used to attach the incoming word into the tree. If no such matching root exists, other phrase structure templates whose terminal elements match this root node are examined in the search for a template that will support attachment. The search is iterated until it is successful (or abandoned, in the case of an ungrammatical string).

For example, consider a tree (Figure 5, panel 1) that has developed to contain a S (sentence) node above a lexically complete NP (noun phrase node, e.g., John). An incoming verb such as saw is then identified as a V (verb) (Panel 2). No template containing V as a terminal element has a root that matches the S-above-NP tree, but the template with VP (verb phrase) as a root permits finding a template with VP as a terminal element and S as its root (Panel 3). These two templates permit attaching V to VP to S (Panel 4). The simple assumption that the processing system selects the first way of attaching the incoming word into the tree that it can find predicts that words will be "minimally attached," i.e., attached into the phrase structure tree with the smallest number of added nodes (see Frazier, 1987a, and Fodor & Inoue, 2000, for further discussion). In the event that two possible attachment operations are equally minimal (requiring the same number of phrase structure templates to connect the incoming word into the existing tree), preference is given to the one that attaches the new word into the phrase currently being processed, presumably because of its greater availability.

Insert Figure 5 about here

Frazier and Rayner (1982) introduced the study of eye movement monitoring into sentence comprehension research, testing some implications of Frazier's (1978) serial, depth-first, garden-path theory. They showed disruption of eye movements in the disambiguating (**boldfaced**) region of sentences like <u>Since Jay always jogs a mile **seems like** a very short distance to him</u> and <u>The second wife will claim the inheritance **belongs to her**</u>. They assumed that the reader had initially constructed the single analysis of the temporarily ambiguous region that was predicted by Frazier's (1978) theory and then had to revise it when disambiguating material indicated it was incorrect. The observed disruption in this case presumably reflected the processing effort required to recognize that the initial analysis was incorrect and to revise it to be consistent with the sentence as it continued.

Support for the predictions of Frazier's minimal attachment and recency ("late closure") principles has been found in a wide range of sentence constructions. Table 1 presents a list of constructions for which reading disruption has been experimentally demonstrated when a temporary ambiguity in the sentence is resolved in favor of the interpretation which is predicted to be unpreferred. However, other theorists have advanced a variety of factors other than structural simplicity that they claimed readers and listeners use in initially parsing a sentence.

Ford, Bresnan, and Kaplan (1982) claimed that, rather than initially constructing the simplest structure for the arguments of a verb, the parser initially selects the structure that is most frequently used for that verb. Abney (1989), Crocker (1985), and Pritchett (1992) developed the claim that the parser made its structural choices by favoring an analysis that treats a phrase as an argument of a verb or other argument assigner, not as an adjunct. Phillips and Gibson (1997) and Stevenson (1994) proposed that the parser prefers to attach a new phrase to the most recent available phrase, not to the phrase which would result in the simplest attachment. Crain and Steedman (1985) and Altmann and Steedman (1988) developed a referential theory of parsing, which claimed that, after projecting all possible interpretations in parallel, a reader or listener initially selects the interpretation which requires the least modifications to the existing discourse structure to permit all its terms to refer successfully.

There is substantial evidence that factors such as frequency, argument structure, recency, and context must be considered in developing a theory of sentence parsing. Frequency of usage of particular structures does seem to have important effects, to be discussed in the following section on lexical factors in comprehension. The argument/adjunct distinction has been the focus of a smaller amount of research, which indicates that a reader or listener does seem to favor analyzing a phrase as an argument rather than an adjunct (Clifton, Speer, & Abney, 1991; Schütze & Gibson, 1999; Speer & Clifton, 1998; note that Clifton et al. provide some evidence that this preference may reflect processes that operate after the creation of an initial structural analysis.) Gibson and his colleagues (e.g., Phillips & Gibson, 1997; Gibson, Pearlmutter, Canseco-Gonsales, & Hickok, 1996) have provided evidence for the importance of recency in several distinct contexts. And a quite substantial amount of research, to be reviewed next, indicates that discourse context affects parsing.

Context effects Some research suggests that discourse context affects parsing in the way described by referential theory (Altmann & Steedman, 1988). This theory claimed that a reader's syntactic knowledge makes available, in parallel, multiple possible interpretations of a sentence. The parser then selects the interpretation whose semantic presuppositions are best satisfied by the discourse context. Most research on the effects of discourse context on parsing has considered phrases that are ambiguous between a modifier usage and some other usage. For instance, consider the sentences <u>The burglar blew open the safe with the dynamite/new lock and made off with the loop</u> (Altmann & Steedman, 1988). In the version containing the phrase <u>with the dynamite</u>, this phrase is most plausibly interpreted a an instrumental modifier of the verb <u>blow open</u> while in the version with <u>with the new lock</u> the phrase modifies the NP <u>the safe</u>. While Frazier's (1987a) garden path theory predicts that the instrument version will be read faster (since the VP attachment of the <u>with</u>-phrase is simpler than the NP attachment; cf. Rayner, Frazier, & Carlson, 1983 for supporting evidence), Altmann & Steedman found that the relative speed of reading the two versions depended on the context. The NP attachment version was read faster than VP attachment in a context that introduced two safes (one with a new lock), thus satisfying the claimed presuppositions of the NP modifier; the VP attachment version was read faster in a context that introduced only a single safe.

Numerous researchers have examined effects like this one (e.g., Altmann, 1988; Altmann, Garnham, & Dennis, 1992; Britt, 1994; Ferreira & Clifton, 1986; Mitchell & Corley, 1994; Murray & Liversedge, 1994; Rayner, Garrod, & Perfetti, 1992; see Mitchell, 1994, for a summary; see Clifton & Ferreira, 1989, and Steedman & Altmann, 1989, for critical discussion of the early research). It does seem clear that a discourse with two possible referents for an unmodified definite NP will disrupt reading, because of the NP's failure to refer successfully. Further, a context that supports a normally-unpreferred modifier interpretation (e.g., the "two safes" context mentioned above) can eliminate the disruption observed for such an interpretation out of context. However, this elimination of garden-pathing seems to be fully effective only when the structural preference for the normally-preferred interpretation is a weak one (e.g. for a prepositional phrase that is an optional argument or an adjunct of the verb, like the instrumental modifier with the dynamite). Garden-pathing is not eliminated when the structural preference is strong (e.g., a preference for a main clause analysis rather than a reduced relative clause analysis as in Bever's (1970) The horse raced past the barn fell, or an adverb's preference to modify a recent vs. a distant verb as in Altmann, van Nice, Garnham, & Henstra's, 1998, She'll implement the plan she proposed next week, or a preference for an obligatory verb argument vs. a NP modifier as in Britt's (1994) He put the book on the battle onto the chair). Rather, in these cases, context that satisfies the presumed presuppositions of a modifier simply reduces the amount of disruption caused by the garden path.

Other types of context may have more substantial effects. Altmann et al. (1998) presented sentences like She'll implement the plan she proposed next week in a context where it answered a question like When will Fiona implement the plan she proposed? This context eliminated or even reversed the normal recency-based preference for the sentence-final adverb to modify the most recent verb. Trueswell and Tanenhaus (1991, 1992) found that the garden-pathing normally observed in a sentence like The student spotted by the proctor was expelled was largely eliminated in a context that described a future event ...tomorrow...a proctor will notice one of the students cheating.

A nonlinguistic context may have even more dramatic effects on sentence comprehension. Tanenhaus and colleagues, (e.g., Sedivy, Tanenhaus, Chambers, & Carlson, 1999; Spivey, Tanenhaus, Eberhard, & Sedivy, 2001; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995) used a head-mounted eyetracker to measure where people looked in an array of objects while they were following auditory instructions about pointing to or manipulating these objects. In this situation, the eyes seem to move quickly to whatever is being referred to (Cooper, 1974). Upon hearing a command like Put the cup on the napkin under the book, the eyes seem to do different things depending on whether the object array contains an empty napkin and two cups (one on a napkin) vs just one cup. The eyes quickly moved to the empty napkin in the one-cup context, suggesting that the on the napkin phrase was taken as a goal, but in the two-cup context, they did not. The latter context may have overridden the default preference to take the on-phrase as an argument of the verb put.

Reanalysis processes. The recognition that such a wide variety of factors affect sentence comprehension has provoked several distinct responses. One response was to turn away from depth-first parsing theories which search for simple principles governing the creation of structural analyses of sentences and develop theories that examine how all sources of information are integrated in choosing a preferred analysis from among all the possible analyses. A second response was to search for grammatical factors other than phrase structure geometries and argument structures that might play a role in the creation of an initial analysis. These two responses will be discussed in later sections. A third response was to maintain the need for a theory of how structural analyses are created but to add additional theoretical claims about how the initial analysis of a sentence is revised in the light of various factors. Rayner, Carlson and Frazier (1983) is an early example of the third type of response. They explored the possibility that semantic and plausibility factors would override initial structural preferences. While they did not find evidence for such an override, they did obtain data that they interpreted to indicate that semantic factors could trigger the revision of an initial syntactic analysis. For instance, they found that a sentence like The spy saw the cop with a revolver exhibited disruption in the prepositional phrase (PP) with a revolver, as if the structurally-simpler analysis in which the PP modifies the verb were initially constructed despite its implausibility and only then modified to the more plausible but more complex analysis in which the PP modifies the noun. They posited the existence of a "thematic processor" that would respond to the existence of a highly plausible alternative to the structurally-preferred analysis.

Such a thematic processor could be developed to explain how a wide range of factors affect the ease of recovering from and revising an initial misanalysis. While no explicit and restricted model exists that successfully explains all the effects of frequency, argument preference, context, etc. in terms of ease of reanalysis, interesting progress toward such a theory is being made (see Fodor & Ferreira, 1998, for a collection of papers exploring theories of reanalysis). A theory of reanalysis must make claims about what kind of information will trigger a revision of a presumed initial analysis, what kind of information guides the revisions that will be made, and what kinds of revisions are easy or difficult. It can claim that some revisions take place so quickly that they are difficult to detect experimentally, while others take substantial time and cause measurable disruption. Fodor and Frazier (1980; cf. Fodor & Inoue, 1994) made a "revision as last resort" claim that seems to have stood up quite well. They proposed that, when readers or listeners are faced with the choice of discarding the grammatical analysis they have built so far and starting over vs. repairing the existing analysis, they choose the latter. Sturt, Pickering, Scheepers, and Crocker (2001) and Schneider and Phillips (2001) have provided experimental support for the claim that revision of an existing analysis is avoided, if possible. They showed (for example) that the phrase beginning <u>had shot</u> in <u>The troops who discovered the enemy spy had shot themselves/himself (and) were later mentioned in the press report</u> is initially attached with <u>The troops</u> as its subject. It is not attached to the more recent phrase <u>enemy spy</u>. From a perspective that claims the human sentence parser has a strong bias toward attaching new phrases to the most recent phrase (Phillips & Gibson, 1997; Stevenson, 1994), this is a surprising result. Presumably, the recency bias is overcome because attaching <u>had shot</u> to the most recent phrase <u>the enemy spy</u> would require revising the initial analysis of <u>the enemy spy</u> as direct object of <u>discovered</u> into an analysis in which it is subject of a sentence complement of <u>discovered</u>.

Fodor and Inoue (2000) questioned the "revision as last resort principle" that Fodor and her colleagues had advanced and that Pickering et al. (2001) and Schneider and Phillips (2001) argued for. Their reasons for doubting the dominance of this principle appealed primarily to data from Japanese and to their intuitive judgments of processing difficulty that seem to conflict with the Pickering et al. and Schneider and Phillips data. Rather than emphasizing the parser's reluctance to revise already-built structure during the reanalysis process, Fodor and Inoue (2000) attempted to interpret evidence for the revision as last resort principle in terms of a very general principle they termed "minimal everything" (and ultimately, "attach quickly"). All choices, both in initial analysis and in reanalysis, are made following a very general least effort principle, formulated in a way that attempts to replace earlier models' explicit reliance on grammatical licensing of the attachments that are made in the parsing and reanalysis processes.

Case structure. Most research on parsing has focused on English, and in English, nearly all information relevant to syntactic structure is carried by word order and by the content of specific lexical items. Modern syntactic theory places substantial emphasis on morphosyntactic features, including case, number, gender, tense, and aspect, and some languages mark these features in ways that could in principle be of substantial use to the parser. Consider case, which many languages mark explicitly. A noun's case (traditionally nominative, accusative, dative, etc.) is related to what role it plays in a verb's argument structure and in turn to its role in sentence structure (subject, object of verb, etc.). While English marks case only on pronouns (where he, she is nominative, her accusative or genitive or dative, etc.), some evidence exist that indicates that this casemarking is used in parsing English (Traxler & Pickering, 1996b). The casemarking on the pronoun she in I knew she and her mother would leave the party early seems to block the normal preference to take a NP after a verb as the direct object of the verb.

Other psycholinguistic researchers have studied case in languages which have much more elaborate casemarking systems (e.g. Kim, 1999, for Korean, and Yamashita, 1997, for Japanese). The most-studied system is German. German determiners are marked for nominative, accusative, dative, and genitive case, but some markings are ambiguous. Further, while the normal word order in German has subject before object, other orders are possible ("scrambling") and motivated in part by pragmatic factors such as focus (see Hemforth & Konieczny, 2000, for an overview). Since German clauses are generally verb-final, it is possible to determine whether the parser assigns case to an NP even before the verb which is the source of the casemarking is encountered. Bader and Meng (1999) have demonstrated that German readers prefer to assign subject before object (nominative before accusative case). Consider the following examples:

a. Die Direktorin hat ersählt, dass die neue Lehrerin einige der Kollegen angerufen hat.

The director has said, that the new teacher some the colleagues phoned had.

("The director said that the new teacher phoned some of the colleagues.")

b. Die Direktorin hat ersählt, dass die neue Lehrerin einige der Kollegen angerufen haben.

The director has said, that the new teacher some the colleagues phoned have.

("The director said that some of the colleagues phoned the new teacher.")

The article die in the embedded clause is ambiguous between nominative and accusative, an ambiguity that is resolved by the number of the final auxiliary verb hat vs haben. Speeded grammaticality judgments were more accurate when the sentence was singular than when it was plural, indicating that the singular NP die neue Leherin had initially been assigned nominative case. Friederici and Mecklinger, 1996, used ERP measurements to reach a similar conclusion.

Bader, Meng, and Bayer (2000) examined the dative case, which is a lexical case needing a specific lexical licenser (while nominative and accusative cases are generally considered structural cases). They provided evidence that structural case is preferred over lexical case in resolving an ambiguity. Their evidence, together with evidence for the existence of preferences among structural cases, led them to claim that the parser builds a substantial amount of syntactic structure, based on its knowledge of possible phrase structure geometries, before reaching the lexical head of a sentence, its main verb.

Prosodic factors in comprehension. Just as most parsing research has been conducted using English, most has used visual presentation of linguistic materials. Fortunately for the field of psycholinguistics, both of these traditions are beginning to change (see, for example, Carreiras, Garcia-Albea, & Sebastian-Galles, 1996; Cutler & Clifton, 2000; De Vincenzi & Lombardo, 2000; Hemforth & Konieczny, 2000; Mazuka & Nagai, 1995). English is just one language among a great many, and it is important not to mistake phenomena specific to English as phenomena that stem from the basic nature of the human sentence parsing mechanism. Even more, speaking and listening are certainly the functions for which our linguistic abilities have evolved. Writing and reading are derived functions, cultural contributions that must, unlike speaking and listening, be explicitly taught.

Warren (1996) and Nicol (1996) contain collections that illustrate some of the recent progress in understanding how prosody can guide comprehension. Much current research is motivated by recent theoretical analyses of prosody, stemming from research by Pierrehumbert (1980; cf. Beckman, 1996; Pierrehumbert& Hirschberg, 1990). These analyses claim that prosody has a structure of its own, constrained by not determined by syntactic structure (Selkirk 1984, 1995). The analyses have been developed into explicit procedures that provide a motivated way of formulating descriptions of the prosody of individual utterances (ToBI, for Tone, Boundary and Intonation; Beckman & Elam, 1997).

Early research focused on demonstrating that prosody can disambiguate utterances (Lehiste, 1973) and on characterizing what kinds of ambiguities can be resolved prosodically (e.g., Nespor & Vogel, 1986). More recent research has examined the question of whether appropriate prosody can overcome the effects of parsing biases and block garden paths. Kjelgaard and Speer (1999; see Marslen-Wilson, Tyler, Warren, Grenier, & Lee, 1992, for a precursor) examined late-closure ambiguities like <u>When Madonna sings the song it's/is a hit</u>. Readers seem to take the phrase the song as the direct object of sings, resulting in a garden path when the sentence continues with the verb is which forces the sentence complement subject analysis of the song. Kjelgaard and Speer recorded sentences like these with either a disambiguating prosodic boundary after the NP the song or after the verb sings. The former prosody is appropriate for the normally-preferred late closure analysis, while the latter is appropriate for an early closure analysis. Kjelgaard and Speer showed that an appropriate prosody eliminated comprehension difficulty of the early closure sentences, relative to a baseline condition whose prosody was appropriate to either interpretation.

Kjelgaard and Speer's claim that prosody determines the initial structuring of an utterance is reinforced by work done by Schafer (1997; Schafer & Speer, 1997). Schafer proposed a "prosodic visibility hypothesis" that suggested that listeners form a prosodic package of material within a single prosodic phrase, and that material within the current prosodic package is more visible than material outside it. This makes attachment within a prosodic package relatively preferred. Thus, when a prosodic phrase ends after angered in The bus driver angered ↓ the rider with a mean look, a listener is more likely to take with a mean look to modify rider than when a prosodic boundary occurs at other points in the utterance.

Schafer, Speer, Warren, and White (2000) demonstrated that these types of prosodic effects could be seen in spontaneously produced speech. They recorded pairs of speakers and listeners talking to one another (using constrained sentence structures) while playing a board game that involved moving pieces from place to place. Their participants used prosodies that effectively disambiguated their utterance and proved to have phonological properties similar to those manipulated in Schafer's earlier work.

Carlson, Clifton and Frazier (2001) developed Schafer's (1997) arguments that the effectiveness of a prosodic boundary depends on the global prosodic representation of a sentence, not just on the local cues. Using a simple auditory questionnaire procedure in which listeners' interpretations of sentences like Susie learned that John telephoned after Bill visited was elicited, Carlson et al. determined that the presence of a prosodic break between John telephoned and after Bill visited could increase the frequency of high attachment interpretations (in which after Bill visited modifies learned), but only when it was prosodically larger than an earlier break after learned. The relative, not the absolute, size of the prosodic boundary seems to guide parsing decisions.

Breadth-first theories of parsing: Constraint-based accounts.

Let us review the material we have covered to this point in our discussion of sentence comprehension. Some of the phenomena point to a process which arrives at a single structural description of a sentence in the fastest possible way, presumably in service of quickly getting some object that can be semantically interpreted. A number of factors have been claimed to guide this process, including structural simplicity (minimal attachment), recency, frequency of use, argument preference, referential success, casemarking, and prosody. Some of these factors (e.g., minimal attachment vs. argument preference vs. referential success) have been presented as mutually-inconsistent alternative accounts of overlapping phenomena. Others can play compatible and complementary roles in a single theory. For instance, a theory that claims that the most quickly constructed syntactically-acceptable analysis is initially accepted can clearly permit the reader's or listener's knowledge of permissible syntactic geometries, casemarking, and prosody, together with the constraints each of these grammatical domains places on the others, to work together in creating a syntactic structure.

However, some of the phenomena we have discussed create difficulties for this view of the parser as a system that follows a depth-first agenda in using a restricted range of grammatical knowledge to create a single, preferred initial analysis of a sentence. The apparent fact that context can fully eliminate at least some structural preferences is accommodated only with great discomfort by a depth-first model, whose only easy response is to say that reanalysis happened too fast to be seen by any existing experimental technique. Further, much development of linguistic theory over the past 20 years has emphasized the specific contributions to lexical information to syntactic structure and de-emphasized rules like phrase-structure rules that specify which particular structural configurations are permissible. These facts, together with a variety of empirical phenomena (some of which will be reviewed below) have encouraged some theorists (including MacDonald et al, 1994, and Tanenhaus & Trueswell, 1995) to develop lexicalist, constraint-based theories of parsing. These theories view parsing as a breadth-first process of weighing the implications that all relevant sources of information (especially lexical sources of information) have for choosing among all possible structural analyses of a sentence.

One major impetus for the development of lexicalist constraint-based theories was research done by Tanenhaus and his colleagues on how different types of lexical information could be combined to influence the interpretation of a sentence (e.g., Carlson & Tanenhaus, 1988; Tanenhaus, Boland, Mauner & Carlson, 1993; Tanenhaus, Garnsey & Boland, 1990). This work was encouraged by the development of connectionist models (cf. Rumelhart & McClelland, 1986), which provided explicit theoretical mechanisms for analyzing the interaction of multiple factors. A major goal of research in the lexicalist constraint-based tradition has been to demonstrate that multiple sources of information interact in determining the preferred resolution of a temporary ambiguity.

A good example of the kind of research that encourages this enterprise is found in Spivey-Knowlton and Sedivy (1995), who measured the self-paced reading time of sentences like The salesman glanced at a/the customer with suspicion/ripped jeans and then walked away. These sentences are temporarily ambiguous in terms of where the PP with suspicion/ripped jeans attaches. Frazier's (1987a) garden-path theory claims that the preferred attachment site is to the verb glance at, predicting slower reading when the content of the PP forces attachment to the NP (i.e., when the PP is with ripped jeans). However, Spivey-Knowlton and Sedivy (1995) showed that the actual preference depended on whether the verb was an action verb (e.g., smash down) vs. a perception verb such as glance at, and whether the object of the verb was definite or indefinite (a vs the customer). Precisely the opposite of the garden-path prediction obtained for perception verbs with an indefinite object, strongly suggesting that a single principle of structural simplicity cannot account for all parsing preferences.

The connectionist metaphor of a network of connected nodes, transmitting activation to one another, underlies much constraint-based lexicalist theorizing (cf. MacDonald et al., 1994). The connectionist emphasis on how connections between nodes are based on experience has encouraged lexicalist theorists to place a good deal of weight on frequency of experience as an important factor in sentence comprehension. Psycholinguistic experimentation has provided a variety of evidence that frequency of different sentence structures does play a role in parsing. For instance, Trueswell (1996) showed that the ease of understanding sentences like The award accepted by the man was very impressive was affected by whether the verb is used more often as a simple past tense verb or as a passive participle (the verb accept has a relatively high frequency of use as a passive participle; a verb like entertain does not). More disruption was observed when the passive participle use required by the sentence was an infrequent usage in the Francis and Kučera (1982) norms. MacDonald (1994) similarly demonstrated more disruption in reading sentences containing a reduced relative clause, such as The rancher could see that the nervous cattle pushed/moved into the crowded pen were afraid of the cowboys, when the verb of the complement sentence was more often used as an intransitive verb (moved) than when it was used most frequently as a transitive (pushed).

Not all researchers find clear and immediate effects of the frequency with which a verb is used in different structures. Ferreira and Henderson (1990) tested an apparently straightforward prediction of a frequency-based account that the difficulty of a sentence containing a temporary ambiguity between a direct object and a sentence complement would be reduced when the verb was most frequently used with sentence complements. They examined eye movements during the reading of sentences like She suspected/pretended Jack owns credit cards and failed to find initial reading time effects of the frequency with which the verb took a direct object vs. a sentence complement; a verb like suspect which infrequently takes a S complement is called an NP-bias verb, while a verb like pretend which quite frequently takes an S complement is called an S-bias verb. In response, Trueswell, Tanenhaus and Kello (1993) argued that Ferreira and Henderson had used implausible NPs as direct objects of their NP-biased verbs, reducing any underlying processing disadvantage for NP-biased verbs. Trueswell et al. successfully demonstrated faster first-pass reading time for sentences like The waiter confirmed/insisted the reservation was made yesterday when the verb was more frequently used with a sentence complement (insisted) than when its most frequent use was with a direct object (<u>confirmed</u>). However, Kennison (2001) countered by saying that Trueswell et al (1993) had themselves confounded plausibility with frequency by using NP complements of S-biased verbs that were less plausible than the NP complements of NP-biased verbs. Kennison gathered eyetracking data from sentences in which plausibility was carefully controlled and found results compatible with those of Ferreira and Henderson.

Some research suggests that readers may not use relative frequency of the type of complement a verb takes to guide their initial parsing decisions but may use it very quickly in arriving at a final analysis. Pickering and Traxler (1998) and Pickering, Traxler and Crocker (2000) demonstrated that reading of sentences like <u>As the woman sailed the magazine about fishing amused all the reporters</u> was disrupted very quickly after the eyes landed on the postverbal noun (<u>magazine</u>) even when the verb was one like <u>hinted</u> that very infrequently is used with a direct object. As discussed earlier (in the section on Immediacy), this finding suggests that the noun is initially taken as direct object of the verb in spite of the low frequency of this analysis and the implausibility of the interpretation of this analysis quickly noted (but cf. Garnsey, Pearlmutter, Myers & Lotocky, 1997, for some apparently-inconsistent evidence; see Gibson, Schütze, & Salomon, 1996 and Gibson & Schütze, 1999 for evidence of the inadequacy of frequency accounts of some other structural preferences; and see Jurafsky, 1996, for a careful and thoroughgoing analysis of how frequency may in fact play the central role in parsing of a wide range of constructions).

In addition to empirical results that suggest frequency, while important, may not be the only important factor in parsing, there are conceptual reasons to question the centrality of frequency in a parsing theory. One reason is that no existing theory adequately specifies how to count frequency, if frequency is treated objectively as the relative frequency with which various constructions have been experienced. In particular, what "grain size" should be counted? The frequency with which particular words co-occur? With which a particular word is used in a specific syntactic construction? With which a class of words is used in a class of equivalent constructions? See Gibson et al. (1996) and Mitchell, Cuetos, Corley and Brysbaert (1995) for further discussion of the grain size problem.

A second reason to question the centrality of frequency is the possibility that production preferences (which determine frequency of experience) and comprehension preferences may be correlated in some instances because they reflect the operation of common underlying factors (e.g., structural simplicity). Showing that comprehension difficulty reflects frequency of experience thus does not mean that one causes the other.

A third reason to be wary of frequency claims is the fact that many experiments demonstrating effects of frequency actually use production measures of frequency. They have subjects complete sentence frames using specified words and tabulate the frequency with which different constructions are used. Merlo (1994) and Roland and Jurafsky (2001) discuss the relative merits of measuring frequency through corpus counts vs. production norms. Carefully-collected production norms may well do the best job of predicting on-line reading effects (cf. Tanenhaus, Spivey-Knowlton and Hanna, 2000, for a clear defense of such norms), but the measures they yield cannot be taken as straightforward measures of the frequency with which particular structures are experienced (and such norms do not in fact always successfully predict on-line comprehension results; Clifton, Kennison, & Albrecht, 1997; Liversedge, Patterson, &

Clayes, 2001).

Even if they are not taken as pure measures of frequency of experience, a constraint-based theorist can argue that production norm frequencies encode the effects of the many lexical factors that should affect comprehension difficulty. Demonstrating that normed properties of particular lexical items, not structural configurations, determine parsing preferences is thus important to the constraint-based lexicalist enterprise. Such demonstrations gain particular value when they are interpreted in explicit, implemented models of parsing decisions, as has been done by Tanenhaus and his colleagues (McRae, Spivey-Knowlton, & Tanenhaus, 1998; Spivey & Tanenhaus, 1998; Tanenhaus et al., 2000). These models are simple networks for choosing between pairs of analyses (e.g., relative clause vs. main clause, as in The defendant/evidence examined by the lawyer turned out to be unreliable) and permit the theorist to compute the explicit consequences of production norm and corpus frequencies together with any other factors the theorist wishes to consider. The models do a good job of fitting reading time measures of sentences that are resolved in favor of the normally-unpreferred analysis, e.g., the reduced relative clause analysis of the sentence just quoted. For instance, they nicely predict the pattern of results obtained by Trueswell, Tanenhaus and Garnsey (1994), who found that using the inanimate subject The evidence in sentences like the one under discussion (a subject that independent norms rated as a good theme but poor agent of the verb examine) eliminated the reading difficulty observed when the subject was The defendant.[1] Some evidence exists, however, that shows that the models fail

[1] Note, Ferreira & Clifton, 1986, who initially measured reading of the ...evidence examined.... sentence and others like it, did not find that inanimacy of the subject eliminated reading disruption. However, Trueswell et al. argued that not all of the Ferreira & Clifton sentences had strong enough lexical biases against the normally-preferred main clause interpretation.

to predict the data from sentences that are resolved in favor of the normally-preferred (main clause) analysis (Binder, Duffy, & Rayner, 2001). The models predict that a context that favors the reduced relative clause interpretation should impair reading of the main clause version, but no such impairment is observed.

Constraint-based theories are not limited to lexical factors. They have readily incorporated evidence (discussed earlier) that discourse context can affect reading difficulty and perhaps guide parsing decisions. Discourse context is treated as just another constraint on parsing (cf. McRae et al., 1998). More than this, constraint-based theories have encouraged experimenters to search for previously-unidentified discourse effects that might affect parsing decisions. For instance, Altmann and Kamide (1999) have used head-mounted eyetracking techniques to determine how quickly stereotyped knowledge of appropriate objects is used in determining the reference of phrases. Their subjects observed a scene on a video display and judged the appropriateness of an auditory sentence describing the scene. The eyes moved faster to a relevant target when the verb in the sentence was thematically appropriate to the target item. For instance, when subjects heard The boy will eat the cake their eyes moved more quickly to a picture of a cake than when they heard The boy will move the cake. Even though this finding does not bear directly on the questions of how ambiguity is resolved that have been the focus of discussion throughout most of this section, it does suggest that listeners very quickly use all their knowledge about specific lexical items in arriving at an interpretation of an utterance.

A final word before concluding this section: Constraint-based lexicalist models of sentence processing have spurred major contributions to our understanding of the topic. They have prompted experimenters to broaden the range of factors they consider in searching for influences on the process of comprehension, and they have prompted theorists to formulate explicit models of how a wide range of information is integrated in choosing among alternative analyses of a sentence. But in doing so, most have given up or sidestepped one of the original goals of parsing theories: to formulate an explicit account of how the parser creates the structure that permits a sentence to be interpreted. As their creators recognize, the models of McRae et al. (1998) and Spivey and Tanenhaus (1998) assume that two (or more) alternative structural analyses of a sentence are available. The models simply describe how the choice between these analyses is made without specifying what makes the analyses available. Other models (e.g., MacDonald et al., 1994) propose that the structure of a clause is projected from its head, the verb. This suggestion is inadequate for a number of reasons (cf. Frazier, 1995), most centrally the fact that in perhaps half the languages of the world the verb is the final element of a clause but there is good reason to doubt that their users wait until the end of clauses to begin parsing sentences.

Future parsing theories will, one hopes, correct the shortcomings of both the depth-first structurally-based theories and the breadth-first constraint-based theories. One possible move would be to assume structural representations that are less dependent on heads of phrases and more amenable to being constructed or activated on the basis of a range of linguistic input. Crocker (1995), Jurafsky (1996) and Vosse and Kempen (2000) can be seen as instances of this approach. Another possible move would be to integrate the depth-first and breadth-first models as they have been developed, permitting more immediate interaction between construction of an initial analysis and its evaluation than has been emphasized in depth-first approaches (cf. Frazier, 1990, as one example of such an attempt).

Long-distance dependencies.

The research discussed to this point has focused on how a reader or listener determines that relations between phrases that are contiguous in a sentence. Consider, however, sentences like What did you read last night and The book I read was by José Saramago and That book was exciting to read. In all these cases, the relation a reader must discover involves the verb read and a distant element (what or the/that book). Discovering this relation poses a substantial challenge to a reader or listener, because the dependency that must be discovered can involve arbitrarily distant elements (e.g., The book that the man from the small town said that he always told his mother that he intended to read was by José Saramago).

Currently dominant linguistic analyses describe sentences like these as having an empty element, a trace, in the position where the "moved" element has to be interpreted (e.g., What is the moved element, and t is the trace in What did you read t last night). Other analyses are possible (see Pickering & Barry, 1991, for a brief survey including psychological implications). Influenced by the dominant view, psycholinguists asked how the human sentence parser discoveres the relation between the moved element, or "filler," and the trace, or "gap." Fodor (1978) analyzed the logical possibilities (e.g., delay assignment of an identified filler to a gap as long as possible) and Frazier (1987b) provided evidence that the preferred strategy was actually to identify the gap position as soon as possible. She presented evidence that Dutch sentences like Karl hielp de mijnwerkers die (t) de boswachter (t) vonden/vond (Karl helped the mineworkers who the forester found plural/singular) were read faster when the number of the final verb was consistent with the early gap (t) position than when it was consistent with the later gap position. That is, extraction from the position of the subject of the final verb was preferred over extraction from object position (cf. Clifton & Frazier, 1989, Fodor, 1989, Stowe, 1986, and Tanenhaus, Boland, Mauner & Carlson, 1993, for discussion of other lines of evidence). DeVincenzi (1991) generalized Frazier's (1987b) "active filler principle" as the "minimal chain principle" (attempt to minimize the lengths of any chain linking a moved element to its trace). She further showed how this principle followed the same underlying logic as minimal attachment and late closure, namely, find the first available analysis.

Tanenhaus et al (1993) argued that a purely syntactic version of the active filler or minimal chain principles is not the whole story. They developed a "thematic filling" hypothesis in which a filler is semantically interpreted as soon as a reader or listener encounters the verb to which it is related, without waiting for the gap position (cf. Pickering & Barry, 1991). They provided evidence for this using a "stops making sense" (incremental unacceptability) task and demonstrated that readers begin rejecting an implausible sentence like Which public library did John contribute some cheap liquor to t last week on the word liquor, before even encountering the word to that licenses the gap position. Results like those of Tanenhaus et al illustrate the point made earlier, that the process of comprehending sentences with long distance dependencies is a complex and subtle one that discourages hope of simplistic models of the sentence comprehension process.

Text and discourse comprehension

The nature of the task.

The core use of a sentence is to convey a message, to express meaningful content. All the parsing mechanisms discussed in the previous section operate in service of this function. Only a small fraction of the research on sentence processing emphasizes how listeners and readers process the meaning of sentences they encounter. The research discussed earlier showing effectively immediate interpretation of sentences constitutes one example, as does research examining how context or preferred thematic structures affects parsing. A few instances of recent research tackle the problems of semantic processing of individual sentences head-on (e.g., Piñango, Zurif & Jackendof, 1999, and McElree, Traxler, Pickering, Seely, & Jackendof, 2001, examining how listeners determine the event structures of predicates, and Frazier, 1999, examining how readers determine the scope of quantifiers). But most psycholinguistic research that has concerned meaning above the level of the word has concentrated on how readers understand texts and connected discourse.

Readers (and listeners) faced with connected discourse have a number of challenging tasks. Among other things, they must determine the reference of referring expressions, especially pronouns and NPs, including determining whether the expression introduces a new entity into the discourse or picks out an existing one. They must determine how any assertion made by one sentence is related to previous assertions. They must (or at least, may) determine what unexpressed assertions follow as implications of what was read. And, according to many researchers, they must create a nonlinguistic representation of the content of what they are reading, a representation that supports the functions of determining reference, relevance, and implications. We will briefly review some core findings about some of these tasks (for discussions of inference, which will not be treated here, see Graesser, Singer & Trabasso, 1994; McKoon & Ratcliff, 1992, 1995; Singer, 1994).

Determination of reference.

Psycholinguists have just begun to examine how a reader or listener picks out the real-world referent of a word or phrase in an ongoing discourse. A promising beginning appears in research mentioned earlier measuring eye movements across an array of objects being talked about (Allopena, Magnuson, & Tanenhaus, 1999; Tanenhaus et al., 1995). These researchers found that the speed of determining the referent of a word (as indexed by how quickly the eyes moved to that referent) is affected by the presence of other objects whose names could be confused with the word denoting the referent. For instance, the speed with which their eyes moved to a candle upon hearing a word beginning can... was slowed by the presence of candy in the array. Similarly, they fixated more rapidly on a tall glass when given the instruction touch the tall glass if the array contained a contrasting object, a short glass. Apparently, listeners interpreted the adjective tall as introducing a contrast with another object in the array and were prepared to interpret a NP as referring to the appropriate member of the contrast.

The bulk of research on how readers and listeners determine reference has focused on anaphoric reference. Pronouns, definite NPs, VP ellipsis, and a variety of other linguistic devices can be used to make reference to an entity or event introduced elsewhere in a discourse, and the use of these devices seems to play an important role in making discourse coherent (Garrod & Sanford, 1994; Grosz & Sidner, 1986). The primary experimental questions that have been asked concern the process of identifying the antecedent of an anaphor and the nature of the representation in which the antecedent is found.

Early research on the comprehension of pronouns emphasized how a reader or listener had to search for the antecedent of an anaphor (Clark & Sengul, 1979; Ehrlich & Rayner, 1983). Later research emphasized how certain antecedents are highly available and how processing is disrupted when a pronoun is forced to refer to a less-available antecedent (e.g., Clifton &

Ferreira, 1987). This approach has been developed most elegantly by Gordon and his colleagues (e.g., Gordon, Grosz, & Gilliom, 1993; Gordon & Scarce, 1995). They analyzed the availability of different potential antecedents in terms of centering theory (Grosz, Joshi, & Weinstein, 1983, 1995), and found that a pronoun whose antecedent occurs in a preceding sentence is comprehended most quickly when its antecedent is a highly-favored "forward-looking center" of that sentence, e.g., its grammatical subject or the first-mentioned entity. Such a pronoun is termed the "backward-looking center" of the sentence that contains it. Gordon and his colleagues (e.g., Gordon & Chan, 1995) also found a reading time penalty for using a repeated name rather than a pronoun (the "repeated name penalty") in the position of the backward looking center, but not in other sentence positions. For instance, following the sentence Susan decided to give Fred a hamster, reading time of Susan told him exactly what do feed it was slowed compared to She told him....

It is tempting to conclude that an appropriate use of a pronoun signals reference to a structurally-defined antecedent, namely, an entity mentioned as topic or center of a preceding sentence. But this overlooks other factors that influence the interpretation of pronouns. For instance, Chambers and Smyth (1998) found that a pronoun in direct object position was most readily interpreted as picking out the direct object of the preceding sentence as its antecedent, and they also found a repeated name penalty for an entity whose antecedent occurred in a syntactically parallel position in the previous sentence. The repeated name penalty disappeared when the name or pronoun occurred in a syntactic position that was not parallel to its antecedent, contrary to centering theory. For example, following Martin Miles told Liz Lovejoy to check the oil, reading of Then Dean Morgan told her/Liz Lovejoy to inspect the coolant was slower when the object was Liz Lovejoy than when it was her. However, this repeated name penalty was not found when the final sentence was Then Dean Morgan told him/Martin Miles to inspect the coolant.

There has been some debate about how quickly the antecedent of a pronoun is determined in reading or listening (see, e.g., Greene, McKoon & Ratcliff, 1992). Ehrlich and Rayner's (1983) eyetracking results have been used to argue that resolution is far from immediate, but their evidence for this conclusion came from sentences in which the antecedent of the pronoun was not readily available, certainly not an appropriate antecedent according to centering theory. In another eyetracking study, Garrod, Freudenthal and Boyle (1994) provided evidence that under some very limited conditions interpretation of a pronoun was essentially immediate. They showed immediate disruption of reading on the verb in sentences like Right away she poured a large glass of coke when the pronoun unambiguously referred to a focused antecedent (i.e., in this case, the female participant had been mentioned in a focused or centered position) and when the verb denoted an incongruent action (the female participant was a passenger and the male steward should have been pouring).

We have been discussing the interpretation of pronouns as if the antecedent of a pronoun is a linguistic object, to be found in the text. This is actually likely to be incorrect (see Garnham, 1999). Garnham argues that the antecedent of an anaphoric expression refers to an element in a mental model, a nonlinguistic conceptual representation. This is an attractive position in that it accounts for a variety of otherwise-puzzling pheneomena. For instance, in Gernsbacher's (1991) "conceptual pronouns," illustrated in I need a plate. Where do you keep them?, the referent of them is clearly not the plate that is needed, but instead is an inferred set of plates that the addressee is assumed to have. Another example is Garnham and Oakhill's (1988) demonstration of the acceptability of reference into anaphoric islands, as in Jim reviewed that book and it will be published in Linguistic Inquiry. The referent of it is Jim's review, an inferred product of the act of reviewing that was described.

The position that at least some anaphors find their referents in a mental model has been extensively examined in the case of VP anaphors. Hankamer and Sag (1976) discuss "deep" vs. "surface" anaphors (termed "interpretive anaphors" and "ellipses" in Sag & Hankamer, 1984). A deep anaphor can be used with or without a linguistic antecedent. Consider a scene where Joe's wife Mary notices that the garbage is, surprisingly, taken out, and Joe says I did it. This interchange is perfectly acceptable. However, it would have been at least a little odd for Joe to have used a surface anaphor and said I did, even though this utterance would have been perfect as a reply to the question Who took out the garbage? There is some evidence that deep and surface VP anaphors are processed differently (Tanenhaus & Carlson, 1990), but there is also evidence that surface form is not irrelevant to the processing of deep VP anaphors (Murphy, 1985). Similarly, in the case of pronouns, which are considered to be deep anaphors, Garnham, Oakhill, Ehrlich and Carreiras (1995) showed faster resolution of the antecedent of pronouns in Spanish and French when gender disambiguation was available, even when the gender of the antecedent was purely formal. This means that the linguistic properties of the antecedent affected comprehension, and suggests that the antecedent is not likely to be purely an object in a nonlinguistic conceptual representation.

Text structure.

Text comprehension became a major subfield of cognitive psychology largely through the efforts of Kintsch (1974; Kintsch & Van Dijk, 1978). Kintsch has proposed a series of models of the process by which the propositions that make up the semantic interpretations of individual sentences are integrated into such larger structures as text bases and situation models. His models described ways in which readers could abstract the main threads of a discourse and infer missing connections, guided by how arguments overlapped across propositions and by linguistic cues signaled by the text and constrained by limitations of short-term store.

Much of Kintsch's work assumed that a reader retrieved information from earlier in the text (information that had been consigned to long-term memory) or from world knowledge only when it was required to interpret a portion of the text which could not be interpreted purely on the basis of local connections. Such long-term memory information contributed to structuring the text only through an effortful and time-consuming process of memory search, triggered by a local coherence break. Current research (including Kintsch's 1988 Construction-Integration model) suggests, in contrast, that a reader or listener regularly makes contact with long-term memory in the process of identifying, evaluating, and expanding on the structure of a discourse.

Some research emphasizes how retrieval of information from long-term memory can be a passive automatic process that occurs automatically throughout comprehension (e.g, McKoon & Ratcliff, 1992, 1998; McKoon, Gerrig & Green, 1996; Myers & O'Brien, 1998). In the Myers and O'Brien "Resonance" model, information in long-term memory is automatically activated by the presence in short-term memory of material that bears an apparently-rough and approximate semantic relation to it. Semantic details, including factors such as negation that drastically change the truth of propositions, do not seem to affect the resonance process. One illustrative piece of evidence for such a process comes from Albrecht and O'Brien (1993). They had subjects read long passages that introduced a character, Mary, as a strict vegetarian. Six sentences after this introduction, a sentence asserted that Mary ordered a cheesburger. Even though this sentence was locally coherent with the preceding few sentences, it was read slowly. O'Brien, Rizella, Albrecht and Halleran (1998) showed the same effect even if the intervening material contained a statement that said Mary was no longer a vegetarian. It appears that the concept of vegetarian was activated by the statement about a cheeseburger, and the apparent inconsistency between the concepts involved slowed reading even if other material that an "intelligent" and strategic process could have used to cancel the conflict had been stated in the text.

Other recent research has emphasized a more active and intelligent search after meaning as the basis by which a reader discovers the conceptual structure of a discourse. Graesser et al. (1994) argued that a reader of a narrative text attempts to build a representation of the causal structure of the text, analyzing events in terms of goals, actions, and reactions. Suh and Trabasso (1993) showed the importance of goals, especially unsatisfied goals, in comprehending a text. If a goal is unsatisfied (e.g., if Jimmy wants a new bicycle, does not receive one as a gift, and then goes to a store with money he has earned), the goal (the bicycle) is more available than if the goal had been satisfied (if Jimmy's mother had given him a bicycle as a gift).

This sensitivity to goal-satisfaction seems to contrast with the crude similarity-based activation that seems to underlie the resonance process discussed earlier. It may be that a resonance process serves as a first stage in processing a text (Rizzella & O'Brien, 1996) and that various reading objectives and details of text structure determine whether or not a reader goes further and searches for a coherent goal structure for the text. Such a perspective at least offers a way of thinking about the peculiar observation (Barton & Sanford, 1993) that readers are perfectly willing to attempt to solve the problem of where to bury the survivors in the context of a description of a plane crash on the French-Spanish border. They do not necessarily probe beyond sheer resonance to recognize that it is inappropriate to bury survivors.

Mental models.

Even if readers and listeners sometimes miss the point, they generally do manage to reach some conclusion about what a text or discourse was about. They construct a model that is not clearly linguistic, a mental model (Johnson-Laird, 1983) or situation model (van Dijk & Kintsch, 1983). The model represents the causal, spatial, temporal and other relations among the events and entities described in the text. We have already introduced the notion of a mental model in the context of comprehending pronouns, advancing Garnham's (1999) suggestion that the referent of a pronoun is found in a mental model. Garnham (1987) provided demonstrations that other NPs find their reference in a mental model, and in fact that discourse had to be understood with respect to a mental model, not just a representation of the form of the discourse. For instance, in the context of a passage that included the sentence By the window was a man with a martini, listeners systematically confused a later sentence from the passage The man standing by the window shouted to the host with The man with the martini shouted to the host in an unexpected memory test.

Some researchers have gone beyond demonstrating the necessity of a mental model and attempted to explore what a mental model contains that a reader uses in comprehension. For example, O'Brien and Albrecht (1992) reported that readers were slowed in reading a sentence that conflicted with an earlier sentence about the location of a protagonist. For instance, in a passage beginning with a statement that located the protagonist, Kim, with respect to a health club, the sentence ....when she saw the instructor come in the door of the club was read faster when Kim had been placed inside the club than when she was outside. Note that this particular manipulation was successful only when readers were instructed to take the perspective of the protagonist. They argued that readers keep track of the spatial location of protagonists in a narrative, information that is naturally represented in a mental model but not in a representation of the language used in the discourse. Glenberg, Meyer, and Lindem (1987) reached a similar conclusion on the basis of probes of entities mentioned in a discourse, showing faster responses to probes of entities that were physically close to the currently-focused protagonist than entities that were further away.

A text can contain cues that help a reader construct a mental model. For instance, connectives such as because, later, meanwhile, next day, ... provide information very relevant to constructing a mental model. Zwaan (1996) showed some apparent processing cost to creating a new time frame in a mental model. In the context of reading a text about Maurice shaking hands at the opening of his new art gallery, a sentence stating that he became very pale was read more slowly when it was introduced by the adverbial connective a day later than by the connective a moment later. Bestgen and Vonk (2000) found that a sentence that shifted topic in a passage (e.g., the French version of the sentence I took my moped from the garage, in the context of cooking dinner in a kitchen) was read more slowly than when the same sentence was read in the context taking a trip in the country. However, in superficial contrast to Zwaan's (1996) result, this topic shift effect disappeared when the topic-shifting sentence began with a temporal adverbial like Around 11 o'clock, serving as a segmentation marker.

One problem in relying on mental models in developing an account of text and discourse comprehension is that existing notions of mental models are rather amorphous, powerful representational devices. This may be unavoidable, if a mental model is taken to be a representation of whatever an individual represents about the world. Psycholinguists may find it useful to develop more constrained theories of mental models, tuned more precisely to the needs of representing discourses. Interesting suggestions about how to proceed exist in the literature on formal semantics (e.g., Barwise & Perry's 1983 Situation Semantics, Heim's 1982, 1983 File Card Semantics and Kamp & Reyle's 1992 Discourse Representation Theory). These theories have in fact stimulated the development of theories of mental models (cf. Garnham, 1999), and one can hope that some of their formal rigor will find its way into theories of how a reader or listener accomplishes the task of mapping speech or text onto meaning.

Summary

In this chapter, we have provided a general overview of language comprehension, from the comprehension of words to sentence and text comprehension. We discussed results of research using both reading and listening (though it is fair to say that there has been more work on the comprehension of written language than spoken language). Along the way, we have mentioned a number of influential models of each of these different aspects of comprehension. There are now fully implemented models of word recognition, both in the visual domain (Coltheart et al., 1993, 2001; Seidenberg & McClelland, 1989) and the auditory domain (McClelland & Elman, 1996; Norris et al., 2000), and there are implemented models of sentence parsing (McRae et al., 1998; Spivey & Tanenhaus, 1998; Vosse & Kempen, 2000) and text comprehension (Kintsch, 1988; Kintsch & van Dijk, 1978; Myers & O'Brien, 1998). There are also implemented models of eye movement control in reading in relation to word processing

(Reichle et al., 1998). What is missing, not only in terms of implemented models, but in terms of general models or theories of language comprehension, are efforts to account for the interaction of word, sentence, and discourse comprehension (and how such activities interact with eye movement control in reading). This will not be an easy task -- see, for example, Just and Carpenter's (1980) attempt to account for these different mental activities in a general model. However, our increasing knowledge of the empirical facts about reading and listening will hopefully make it possible that an all-encompassing model of language comprehension will be developed in the near future.

References

Abney, S. (1989). A computational model of human parsing. Journal of Psycholinguistic Research, 18, 129-144.

Albrecht, J. E., & O'Brien, E. J. (1993). Updating a mental model: Maintaining both local and global coherence. Journal of Experimental Psychology: Learning, Memory, and Cognition, 19, 1061-1070.

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

Altmann, G. (1988). Ambiguity, parsing strategies, and computational models. Language and Cognitive Processes, 3, 73-98.

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.

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

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

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

Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. Psychonomic Bulletin & Review, 4, 439-461.

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). The CELEX lexical database (CD-ROM). Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

Bader, M., & Meng, M. (1999). Subject-object ambiguities in German embedded clauses: An across-the-board comparison. Journal of Psycholinguistic Research, 28, 121-144.

Balota, D.A. (1983). Automatic semantic activation and episodic memory encoding. Journal of Verbal Learning and Verbal Behavior, 22, 88-104.

Balota, D.A., & Chumbley, J.I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. Journal of Experimental Psychology: Human Perception and Performance, 10, 340-357.

Balota, D.A., & Chumbley, J.I. (1985). The locus of word-frequency effects in the pronunciation task: Lexical access and/or production? Journal of Memory and Language, 24, 89-106.

Balota, D.A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. Cognitive Psychology, 17, 364-390.

Baron, J., & Strawson, C. (1976). Use of orthographic and word-specific knowledge in reading words aloud. Journal of Experimental Psychology: Human Perception and Performance, 2, 386-393.

Barton, S. B., & Sanford, A. J. (1993). A case study of anomaly detection: Shallow semantic processing and cohesion establishment. Memory & Cognition, 21, 477-487.

Barwise, J., & Perry, J. (1983). Situations and attitudes. Cambridge, MA: MIT Press.

Bauer, D., & Stanovich, K.E. (1980). Lexical access and the spelling-to-sound regularity effect. Memory & Cognition, 8, 424-432.

Beckman, M. (1996). The parsing of prosody. Language and Cognitive Processes, 11, 17-67.

Beckman, M. E., & Elam, G. A. (1997). Guidelines for ToBI labelling, version 3.0. Manuscript and accompanying speech materials.: Ohio State University.

Berent, I., & Perfetti, C.A. (1995). A rose is a reez: The two cycles model of phonology assembly in reading English. Psychological Review, 102, 146-184.

Besner, D., Twilley, L., McCann, R., & Seergobin, K. (1990). On the connection between connectionism and data: Are a few words necessary? Psychological Review, 97, 432-446.

Bestgen, Y., & Vonk, W. (2000). Temporal adverbials as segmentation markers in discourse comprehension. Journal of Memory and Language, 42, 74-87.

Bever, T. G. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), Cognition and the development of language (pp. 279-352). New York: Wiley.

Bever, T. G., Sanz, M., & Townsend, D. J. (1998). The emperor's psycholinguistics. Journal of Psycholinguistic Research, 27, 261-284.

Binder, K., Duffy, S., & Rayner, K. (2001). The effects of thematic fit and discourse context on syntactic ambiguity resolution. Journal of Memory and Language, in press

Boland, J.E., Tanenhaus, M.K., & Garnsey, S.M. (1990). Evidence for the immediate use of verb control information in sentence processing. Journal of Memory and Language, 29, 413-432.

Britt, M. A. (1994). The interaction of referential ambiguity and argument structure in the parsing of prepositional phrases. Journal of Memory and Language, 33, 251-283.

Buckner, R. L. (1998). Event-related fMRI and the hemodynamic response. Human Brain Mapping, 6, 373-377.

Burgess, C., Seidenberg, M.S., & Tanenhaus, M.K. (1989). Context and lexical access: Implications of nonword interference for lexical ambiguity resolution. Journal of Experimental Psychology: Learning, Memory, and Cognition, 15, 620-632.

Carlson, G. N., & Tanenhaus, M. K. (1988). Thematic roles and language comprehension. In W. Wilkins (Ed.), Syntax and Semantics: Thematic relations (pp. 263-300). New York: Academic Press.

Carlson, K., Clifton, C., Jr., & Frazier, L. (2001). Prosodic boundaries in adjunct attachment. Journal of Memory and Language.

Carreiras, M., Garcia-Albea, J., & Sebastian-Galles, N. (1996). Language processing in Spanish.

Mahwah, NJ: Elrbaum.

Cattell, J.M. (1886). The time it takes to see and name objects. Mind, 11, 63-65.

Chambers, C. G., & Smyth, R. (1998). Structural parallelism and discourse coherence: A test of centering theory. Journal of Memory and Language, 39, 593-608.

Chumbley, J.I., & Balota, D.A. (1984). A word's meaning affects the decision in lexical decision. Memory & Cognition, 12, 590-606.

Church, K., & Patil, R. (1982). Coping with syntactic ambiguity, or how to put the block in the box on the table. American Journal of Computational Linguistics, 8, 139-149.

Clark, H. H., & Clark, E. V. (1977). Psychology and language. NY: Harcourt Brace Jovanovich.

Clark, H. H., & Sengul, C. J. (1979). In search of referents for nouns and pronouns. Memory & Cognition, 7, 35-41.

Clifton, C., Jr. (1993). Thematic roles in sentence parsing. Canadian Journal of Psychology, 47, 222-246.

Clifton, C., Jr., & Duffy, S.A. (2001). Sentence and text comprehension: Roles of linguistic structure. Annual Review of Psychology, in press.

Clifton, C., Jr., & Ferreira, F. (1987). Discourse structure and anaphora: Some experimental results. In M. Coltheart (Ed.), Attention and performance 12: The psychology of reading. (pp. 635-654): Lawrence Erlbaum Associates, Inc, Hove, England.

Clifton, C., Jr., & Ferreira, F. (1989). Ambiguity in context. Language and Cognitive Processes, 4, SI 77-104.

Clifton, C., Jr., Kennison, S., & Albrecht, J. (1997). Reading the words her, his, him: Implications for parsing principles based on frequency and structure. Journal of Memory and Language, 36, 276-292.

Clifton, C., Jr., & Odom, P. (1966). Similarity relations among certain English sentence constructions. Psychological Monographs, 80, No. 5 (Whole No. 613).

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.

Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), Strategies of information processing (pp 151-216). London:Academic Press.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed processing. Psychological Review, 100, 589-608.

Coltheart, M, Patterson, K., & Marshall, J.C. (1980). Deep dyslexia. London: Routledge.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recogntion and reading aloud. Psychological Review.

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. Cognitive Psychology, 6, 61-83.

Crocker, M. (1995). Computational psycholinguistics: An interdisciplinary approach to the study of language. Dordrecht: Kluwer.

Cutler, A., & Clifton, C., Jr. (2000). Blueprint of the listener. In P. Hagoort & C. Brown (Eds.), The neurocognition of language (pp. 123-166). Oxford: Oxford University Press.

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

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. Kartunnen & A. Zwicky (Eds.), Natural language parsing (pp. 320-358): Cambridge: Cambridge University Press.

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..M., & Davidson, M. (1995). Time course of phonological activation during reading: Evidence from eye fixations. Journal of Experimental Psychology: Learning, Memory, and Cognition, 17, 618-632.

DeVincenzi, M. (1991). Syntactic parsing strategies in Italian. Dordrecht: Kluwer.

De Vincenzi, M., & Lombardo, V. (Eds.). (2000). Cross-linguistic perspectives on language processing. Dordrecht: Kluwer.

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., Kambe, G., & Rayner, K. (2001). The effect of prior disambiguating context on the comprehension of ambiguous words: Evidence from eye movements. In D. Gorfein (Ed.), On the consequences of meaning selection. American Psychological Association Press.

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.

Ellis, A.W., & Ralph, M.A. L. (2000). Age of acquisition effects in adult lexical processing reflect loss of plasticity in maturing systems: Insights from connectionist networks. Journal of Experimental Psychology: Learning, Memory, and Cognition, 26, 1103-1123.

Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. Journal of Memory & Language, 25(3), 348-368.

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

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

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

Fodor, J. A., Bever, T. G., & Garrett, M. F. (1974). The psychology of language: An introduction to psycholinguistics and generative grammar. New York:McGraw-Hill.

Fodor, J. D. (1978). Parsing strategies and constraints on transformations. Linguistic Inquiry, 9, 427-474.

Fodor, J. D. (1989). Empty categories in sentence processing. Language and Cognitive Processes, 4, SI 155-210.

Fodor, J. D., & Ferreira, F. (Eds.). (1998). Sentence reanalysis. Dordrecht: Kluwer

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

Fodor, J. D., & Inoue, A. (2000). Garden path re-analysis: Attach (anyway) and revision as last resort. In M. D. Vincenzi & V. Lombardo (Eds.), Cross-linguistic perspectives in language processing. Dordrecht: Kluwer

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.

Ford, M., Bresnan, J., & Kaplan, R. (1982). A competence-based theory of syntactic closure. In

J. Bresnan (Ed.), The mental representation of grammatical relations.: Cambridge, Mass: The MIT Press, pp 727-796.

Forster, K. I. (1979). Levels of processing and the structure of the language processor. In W. E. Cooper & E. C. T. Walker (Eds.), Sentence Processing (pp. 27-86).Hillsdale, NJ: Erlbaum.

Forster, K.I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. Journal of Experimental Psychology: Learning, Memory, and Cognition, 10, 680-698.

Foss, D. J., & Blank, M.A. (1980). Identifying the speech code. Cognitive Psychology, 12, 1-31.

Foss, D.J., & Jenkins, C.M. (1973). Some effects of context on the comprehension of ambiguous sentences. Journal of Verbal Learning and Verbal Behavior, 12, 577-589.

Francis, W. N., & Kučera, H. (1982). Frequency analysis of English usage. Houghton-Mifflin.

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

Frazier, L. (1983). Processing sentence structure. In K. Rayner (Ed.), Eye movements in reading: Perceptual and language processes (pp. 215-236). New York: Academic Press.

Frazier, L. (1987a). Sentence processing: A tutorial review. In M. Coltheart (Ed.), Attention and performance (pp. 559-586). Hillsdale, NJ: Lawrence Erlbaum Associates.

Frazier, L. (1987b). Syntactic processing: Evidence from Dutch. Natural Language and Linguistic Theory, 5, 519-559.

Frazier, L. (1990). Exploring the architecture of the language system. In G. Altmann (Ed.), Cognitive models of speech processing: Psycholinguistic and computational perspectives (pp. 409-433). Cambridge, MA: MIT Press.

Frazier, L. (1995a). Representational issues in psycholinguistics. In J. L. Miller & P. D. Eimas (Eds.), Handbook of perception and cognition, volume 11: Speech, language, and communication. New

York: Academic Press.

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

Frazier, L. (1999). On sentence interpretation. Dordrecht: Kluwer

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

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

Frazier,L., & Rayner, K (1987) Resolution of syntactic category ambiguities: Eye movements 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.

Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and false trails. Psychological Bulletin, 123, 71-100.

Garnham, A. (1981). Mental models as representations of text. Memory & Cognition, 6, 560-565.

Garnham, A. (1999). Reference and anaphora. In S. Garrod & M. J. Pickering (Eds.), Language processing (pp. 335-362). Hove, UK: Psychology Press.

Garnham, A., & Oakhill, J. (1988). "Anaphoric islands" revisited. Quarterly Journal of Experimental Psychology, 40A, 719-735.

Garnham, A., Oakhill, J., Ehrlich, M.-F., & Carreiras, M. (1995). Representations and processes in the interpretation of pronouns: New evidence from Spanish and French. Journal of Memory and Language, 34, 41-62.

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.

Garrett, M. F. (1976). Syntactic processes in sentence production. In R. J. Wales & E. Walker (Eds.), New approaches to language mechanisms. A collection of psycholinguistic studies.: Amsterdam: North-Holland.

Garrod, S., Freudenthal, D., & 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., & Pickering, M. (Eds.). (1999). Language Processing. Hove: Psychology Press.

Garrod, S., & Sanford, A. J. (1994). Resolving sentences in a discourse context: How discourse representation affects language understanding. In M. A. Gernsbacher (Ed.), Handbook of Psycholinguistics. San Diego: Academic Press.

Gazzaniga, M.S. (2000). Cognitive neuroscience: A reader. Oxford: Blackwell.

Gernsbacher, M.A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. Journal of Experimental Psychology: General, 113, 256-280.

Gernsbacher, M.A. (1991). Comprehending conceptual anaphors. Language and Cognitive Processes, 6, 81-105.

Gernsbacher, M.A. (Ed.) (1994). Handbook of Psycholinguistics. San Diego: Academic Press.

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

Gibson, E., Pearlmuter, N., Canseco-Gonzalez, E., & Hickok, G. (1996). Recency preference in the human sentence processing mechanism. Cognition, 59, 23-59.

Gibson, E., & Schütze, C. T. (1999). Disambiguation preferences in noun phrase conjunction do not mirror corpus frequency. Journal of Memory and Language, 40, 263-279.

Gibson, E., Schütze, C. T., & Salomon, A. (1996). The relationship between the frequency and the processing complexity of linguistic structure. Journal of Psycholinguistic Research, 25, 59-92.

Goodman, K.S. (1970). Reading: A psycholinguistic guessing game. In H. Singer and R.B. Ruddell (Eds.), Theoretical models and processes of reading. Newark: International Reading Association.

Gordon, P. C., & Chan, D. (1995). Pronouns, passives, and discourse coherence. Journal of Memory and Language, 34, 216-231.

Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, nouns, and the centering of attention in discourse. Cognitive Science, 17, 311-349.

Gordon, P. C., & Scearce, K. A. (1995). Pronominalization and discourse coherence, discourse structure and pronoun interpretation. Memory & Cognition, 23, 313-323.

Gough, P. B. (1965). Grammatical transformations and speed of understanding. Journal of Verbal Learning and Verbal Behavior, 4, 107-111.

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

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. Psychological Review, 101, 371-395.

Greene, S. B., McKoon, G., & Ratcliff, R. (1992). Pronoun resolution and discourse models. Journal of Experimental Psychology: Learning, Memory, and Cognition, 18, 266-283.

Grosz, B., Joshi, A., & Weinstein, S. (1983). Providing a unified account of definite noun phrases in discourse. Proceedings of the Association for Computational Linguistics, MIT, 44-50.

Grosz, B. J., Joshi, A. K., & Weinstein, S. (1995). Centering: a framework for modeling the local coherence of discourse. Computational Linguistics, 21, 203-225.

Grosz, B., & Sidner, C. (1986). Attention, intention, and the structure of discourse. Computational Linguistics, 12, 175-204.

Haberlandt, K. (1994). Methods in reading research. In M.A. Gernsbacher (Ed.), Handbook of

Psycholinguistics (pp 1-32). San Diego, Ca: Academic Press.

Hankamer, J., & Sag, I.(1976).Deep and surface anaphora. Linguistic Inquiry, 7, 391-426.

Harm, M.W., & Seidenberg, M.S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. Psychological Review, 106, 491-528.

Heim, I. (1982). The semantics of definite and indefinite noun phrases, Dissertation, UMass.

Heim, I. (1983). File change semantics and the familiarity theory of definiteness. In R. Bauerle, C. Schwarze, & A. von Stechow (Eds.), Meaning, Use, and the Interpretation of Language. Berlin:de Gruyter.

Hemforth, B., & Konieczny, L. (2000). German sentence processing. Dordrecht: Kluwer.

Hemforth, B., & Konieczny, L. (2000). Cognitive parsing in German: An introduction. In B. Hemforth & L. Konieczny (Eds.), German sentence processing. Dordrecht: Kluwer.

Henderson, J.M. and 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.

Hochberg, J. (1970). Components of literacy: Speculations and exploratory research. In H. Levin and J.P. Williams (Eds.), Basic studies on reading. New York: Basic Books.

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

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-Laird, P. N. (1983). Mental models: Cambridge, MA: Harvard Univ. Press.

Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. Cognitive Science, 20, 137-194.

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

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

Kamp, H., & Reyle, U. (1992). From discourse to logic. Dordrecht: Kluwer.

Kennison, S. M. (2001). Limitations on the use of verb information during sentence comprehension. Psychonomic Bulletin & Review, in press

Kennison, S.M. and 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.

Kim, J. J.(1997).Children's theory of race: a question of interpretation. Cognition, 64, 345.

Kimball, J. (1973). Seven principles of surface structure parsing in natural language. Cognition, 2, 15-47.

Kintsch, W. (1974). The representation of meaning in memory and knowledge.: New York:Wiley.

Kintsch, W. (1988). The use of knowledge in discourse processing. Psychological Review, 95, 163-182.

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

Kutas, M., & Van Petten, C.K. (1994). Psycholinguistics electrified: Event-related brain potential investigations. In M.A. Gernsbacher (Ed.), Handbook of Psycholinguistics (pp 83-144). Academic Press.

Lee, H., Rayner, K., & Pollatsek, A. (1999). The time course of phonological, semantic, and orthographic coding in reading: Evidence from the fast priming technique. Psychonomic Bulletin & Review, 6, 624-634.

Lee, H., Rayner, K., & Pollatsek, A. (2001). The relative contribution of consonants and vowels to word identification in reading. Journal of Memory and Language, in press.

Lee, Y., Binder, K.S., Kim, J., Pollatsek, A., & Rayner, K. (1999). Activation of phonological codes during eye fixations in reading. Journal of Experimental Psychology: Human Perception and Performance, 25, 948-964.

Lehiste, I. (1973). Phonetic disambiguation of syntactic ambiguity. Glossa, 7, 107-122.

Lesch, M.F., & Pollatsek, A.(1998). Evidence for the use of assembled phonology in accessing the meaning of printed wordsJournal of Experimental Psychology:Learning,Memory,andCognition,24,573-592.

Levin, H., & Kaplan, E.L. (1970). Grammatical structure and reading. In H. Levin and J.P. Williams (Eds.), Basic studies on reading. New York: Basic Books.

Lewellen, M.J., Goldinger, S.D., Pisoni, D.B., & Greene, B. (1993). Lexical familiarity and processing efficiency: Individual differences in naming, lexical decision, and semantic categorization. Journal of Experimental Psychology: General, 122, 316-330.

Lewis, M.B., Gerhand, S., & Ellis, H.D. (2001). Re-evaluating age-of-acquisition effects: Are they simply cumulative-frequency effects? Cognition, 78, 189-205.

Liversedge, S. P., Paterson, K. B., & Clayes, E. L. (2001). The influence of only on syntactic processing of "long" relative clause sentences. Quarterly Journal of Experimental Psychology, in press.

Lukatela, G., & Turvey, M.T. (1998). Reading in two alphabets. American Psychologist, 53, 1057-1077.

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

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. Psychological Review, *101*, 676-703.

Marcel, A. J. (1983). Conscious and unconscious perception: Experiments on visual masking and word recognition. Cognitive Psychology, 15, 197-237.

Marcus, M.(1980). A theory of syntactic recognition for natural language. Cambridge: MIT Press.

Marslen-Wilson, W. D. (1973). Linguistic structure and speech shadowing at very short latencies. Nature, 244, 522-523.

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

Marlsen-Wilson, W. D., Tyler, L.K., Waksler, R., & Obler, L. (1994). Morphology and meaning in the English mental lexicon. Psychological Review, 101, 3-33.

Marslen-Wilson, W. D., Tyler, L. K., Warren, P., Grenier, P., & Lee, C. S. (1992). Prosodic effects in minimal attachment. Quarterly Journal of Experimental Psychology,45A,73-87.

Marslen-Wilson, W.D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. Cognitive Psychology, 10, 29-63.

Masson, M.E.J. (1983). Conceptual processing of text during skimming and rapid sequential reading. Memory & Cognition, 11, 262-274.

Mazuka, R., & Nagai, N. (1995). Japanese Sentence Processing. Hillsdale, NJ: Erlbaum.

McClelland, J.L., & Elman, J.L. (1986). The TRACE model of speech perception. Cognitive Psychology, 18, 1-86.

McClelland, J.L., & Rumelhart, D.E. (1981). An interactive activation model of context effects in letter perception: Part 1.An account of basic findings. Psychological Review,88,375-407.

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

McElree, B. (1993). The locus of lexical preference effects in sentence comprehension: A time-course analysis. Journal of Memory and Language, 32, 536-571.

McElree, B., Traxler, M., Pickering, M. J., Seely, R. E., & Jackendoff, R. (2001). Reading time evidence for enriched composition. Cognition, 78, B17-B25.

McKoon, G., Gerrig, R. J., & Greene, S. B. (1996). Pronoun resolution without pronouns: Some consequences of memory based text processing. Journal of Experimental Psychology: Learning, Memory, and Cognition.

McKoon, G., & Ratcliff, R. (1992). Inference during reading. Psychological Review, 99, 440-466.

McKoon, G., & Ratcliff, R. (1995). The minimalist hypothesis: Directions for research. In G. Weaver, S. Mannes, & C. R. Fletcher (Eds.), Discourse comprehension. Hillsdale, NJ: Erlbaum.

McKoon, G., & Ratcliff, R. (1998). Memory-based language processing: Psycholinguistic research in the 1990s. Annual Revieww of Psychology, 49, 25-42.

McQueen, J.M., & Cutler, A. (1998). Morphology in word recognition. In A. Spencer & A.M. Zwicky (Eds.), The handbook of morphology (pp 406-427). Oxford: Blackwell.

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.

Mehler, J. (1963). Some effects of grammatical transformations on the recall of English sentences. Journal of Verbal Learning and Verbal Behavior, 2, 346-351.

Mehler, J., Sequi, J., & Carey, P. (1978). Tails of words: Monitoring ambiguity. Journal of Verbal Learning and Verbal Behavior, 17, 29-36.

Merlo, P. (1994). A corpus-based analysis of verb continuation frequencies for syntactic processing. Journal of Psycholinguistic Research, 23, 435-457.

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.

Meyer, D.E., Schvaneveldt, R.W., & Ruddy, M.G. (1974). Functions of graphemic and phonemic codes in visual word-recognition. Memory & Cognition, 2, 309-321.

Mitchell, D. C. (1994). Sentence parsing. In M. A. Gernsbacher (Ed.), Handbook of psycholinguistics (pp. 375-410). New York: Academic Press.

Mitchell, D. C., & Corley, M. M. B. (1994). Immediate biases in parsing: Discourse effects or experimental artifacts?Journal of Experimental Psychology: Learning, Memory, and Cognition,20,217-222.

Mitchell, D. C., Cuetos, F., Corley, M. M. B., & Brysbaert, M. (1995). Exposure-based models of human parsing: Evidence for the use of coarse-grained (non-lexical) statistical records. Journal of Psycholinguistic Research, 24, 469-488.

Monsell. S., Doyle, M.C., & Haggard, P.N. (1989). Effects of frequency on visual word recognition tasks: Where are they? Journal of Experimental Psychology: General, 118. 43-71.

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.

Morrison, C.M., & Ellis, A.W. (1995). Roles of word frequency and age of acquisition in word naming and lexical decision. Journal of Experimental Psychology: Learning, Memory, and Cognition, 21, 116-133.

Morton, J. (1964). The effect of context on the visual duration thresholds for words. British Journal of Psychology, 55, 165-180.

Murphy, G. L. (1985). Processes of understanding anaphora. Journal of Memory and Language, 24, 290-303.

Murray, W. S.,& Liversedge, S. P.(1994).Referential context effects on syntactic processing. In C. Clifton, L.Frazier &K.Rayner (Eds.),Perspectives on sentence processing (pp.359-388). Hillsdale: Erlbaum

Myers, J. L., & O'Brien, E. J. (1999). Accessing the discourse representation while reading. Discourse Processes, 26, 131-157.

Nespor, M., & Vogel, I. (1986). Prosodic phonology. Dordrecht: Foris.

Newman, J.E., & Dell, G.S. (1978). The phonological nature of phoneme monitoring: A critique of some ambiguity studies. Journal of Verbal Learning and Verbal Behavior, 17, 359-374.

Norris, D., McQueen, J.M., & Cutler, A. (2000). Merging information in speech recognition: feedback is never necessary. Behavioral and Brain Sciences, 23, 299-370.

O'Brien, E. J., & Albrecht, J. E. (1992). Comprehension strategies in the development of a mental model. Journal of Experimental Psychology: Learning, Memory and Cognition,18,777-784.

Onifer, W., & Swinney, D.A. (1981). Accessing lexical ambiguities during sentence comprehension: Effects of frequency of meaning and contextual bias. Memory & Cognition, 15, 225-236.

Paap, K.R.,Newsome,S.L.,McDonald,J.E.,&Schvaneveldt,R.W.(1982). An activation-verification model for letter and word recognition: The word superiority effect. Psychological Review, 89, 573-594.

Perfetti, C.A., & Hogaboam, T.W. (1975). The relationship between single word decoding and reading comprehension skill. Journal of Educational Psychology, 67, 461-469.

Petersen, S.E., Fox, P.T., Posner, M.I., Mintun, M., & Raichle, M.E. (1989). Positron emission tomographic studies of the processing of single words. Journal of Cognitive Neuroscience, 1, 153-170.

Phillips, C., & Gibson, E. (1997). On the strength of the local attachment preference. Journal of

Psycholinguistic research, 26(3), 323-346.

Pickering, M., & Barry, G. (1991). Sentence processing without empty categories. Language and Cognitive Processes, 6, 229-259.

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

Pickering, M. W., & Traxler, M. (2000). Parsing and incremental understanding during reading. In M. W. Crocker & M. Pickering & J. C. Clifton (Eds.), Architectures and mechanisms for language processing (pp. 238-258). Cambrige: Cambridge University Press.

Pickering, M. J., Traxler, M. J., & Crocker, M. W. (2000). Ambiguity resolution in sentence processing: Evidence against likelihood. Journal of Memory and Language, 43, 447-475.

Pierrehumbert, J. B. (1980). The phonology and phonetics of English intonation. PhD thesis, MIT.

Pierrehumbert, J. B., & Hirschberg, J. (1990). The meaning of intonational contours in the interpretation of discourse. In P. R. Cohen & J. Morgan & M. E. Pollack (Eds.), Intentions in communication (pp. 271-311). Cambridge, MA: MIT Press.

Piñango, M. M., Zurif, E., & Jackendoff, R. (1999). Real-time processing implications of enriched composition at the syntax-semantics interface. Journal of Psycholinguistic Research, 28, 395-414.

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

Pollatsek, A., Lesch, M., Morris, R.M., & 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.

Potter, M.C., Kroll, J.F., & Harris, C. (1980). Comprehension and memory in rapid, sequential reading. In R.S. Nickerson (Ed.), Attention and performance (vol 8). Hillsdale, NJ: Erlbaum.

Pritchett, B. L. (1992).Grammatical competence and parsing performance. Univ of Chicago Press.

Raney, G.E., & Rayner, K. (1993). Event-related brain potentials, eye movements, and reading. Psychological Science, 4, 283-286.

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

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

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

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

Rayner, K., Binder, K.S., & Duffy, S.A. (1999). Contextual strength and the subordinate bias effect: Comment on Martin, Vu, Kellas, and Metcalfe. Quarterly Journal of Experimental Psychology, 52A, 841-852.

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

Rayner, K., Garrod, S.C., & Perfetti, C.A. (1992). Discourse influences during parsing are delayed. Cognition, 45, 109-139.

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

Rayner, K., & Sereno, S.C. (1994). Eye movements in reading: Psycholinguistic studies. In M.A. Gernsbacher (Ed.), Handbook of Psycholinguistics (pp 57-82). San Diego: Academic Press.

Rayner, K., Sereno, S.C., Lesch, M.F., & Pollatsek, A. (1995). Phonological codes are automatically activated during reading: Evidence from an eye movement priming paradigm. Psychological Science, 6, 26-30.

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.

Reicher, G.M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. Journal of Experimental Psychology, 81, 275-280.

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.

Riesbeck, C. K., & Schank, R. (1978). Comprehension by computer: Expectation-based analysis of sentences in context. In W. J. M. Levelt & G. B. F. d'Arcais (Eds.), Studies in the perception of language (pp. 247-294). New York: Wiley.

Rizzella, M. L., & O'Brien, E. J. (1996). Accessing global causes during reading. Journal of

Experimental Psychology: Learning, Memory, and Cognition, 22, 1208-1218.

Roland, D., & Jurafsky, D. (2001). Verb sense and verb subcatorization probabilities. In S. Stevenson & P. Merlo (Eds.), Papers from the 1998 CUNY sentence processing conference.: Benjamins.

Rumelhart, D., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: Part 2. The contextual enhancement effect and some tests and extensions of the model. Psychological Review, 89, 60-94.

Rumelhart, D., & McClelland, J. L. (Eds.). (1986). Parallel distributed processing: Explorations in the microstructure of cognition.: (Vol 1). Cambridge, MA: MIT Press.

Sag, I., & Hankamer, J. (1983). Toward a theory of anaphoric processing. Linguistics and Philosophy, 5, 325-345.

Samuel, A. (1981). Phonemic restoration: Insights from a new methodology. Journal of Experimental Psychology, 110, 474-494.

Samuel, A. (1996). Phoneme restoration. Language and Cognitive Processes, 11,647-654.

Schafer, A. (1997). Prosodic parsing: The role of prosody in sentence comprehension. Unpublished PhD Dissertation, University of Massachusetts, Amherst, MA.

Schafer, A. J., & Speer, S. R. (1997). The role of prosodic phrasing in sentence comprehension. Paper presented at CUNY Conference on Human Sentence Processing, March, 1997, Santa Monica, CA.

Schafer, A. J., Speer, S. R., Warren, P., & White, S. D. (2000). Intonational disambiguation in sentence production and comprehension. Journal of Psycholinguistic Research, 29, 169-182.

Schilling, H.E.H., Rayner, K., & Chumbley, J.I. (1998). Comparing naming, lexical decision, and eye fixation times:Word frequency effects and individual differences. Memory & Cognition, 26,1270-1281.

Schneider, D., & Phillips, C. (2001). Grammatical search and reanalysis. Journal of Memory and

Language, in press

Schuberth,R.E.,& Eimas, P.D.(1977). Effects of context on the classification of words and nonwords. Journal of Experimental Psychology: Human Perception and Performance, 3, 27-36.

Schütze, C. T., & Gibson, E. (1999). Argumenthood and English prepositional phrase attachment. Journal of Memory and Language, 40, 409-431.

Sedivy, J. C., Tanenhaus, M., K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. Cognition, 71, 109-147.

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

Seidenberg, M.S., Tanenhaus, M.K., Leiman, J.M., Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. Cognitive Psychology, 14, 489-537.

Selkirk, E. (1984). Phonology and syntax: The relation between sound and structure. MIT Press.

Selkirk, E. O. (1995). Sentence prosody: Intonation, stress, and phasing. In J. Goldsmith (Ed.), Handbook of phonological theory (pp. 550-569). Oxford: Blackwell.

Sereno, S.C., & Rayner, K. (1992). Fast priming during eye fixations in reading. Journal of Experimental Psychology: Human Perception and Performance, 18, 173-184.

Sereno, S.C., & Rayner, K. (2000a). The when and where of reading in the brain. Brain and Cognition, 42, 78-81.

Sereno, S.C., & Rayner, K. (2001b). Spelling-sound regularity effects on eye fixations in reading. Perception & Psychophysics, 62, 402-409.

Sereno, S.C., Rayner, K., & Posner, M.I. (1998). Establishing a time-line of word recognition: Evidence from eye movements and event-related potentials. NeuroReport, 9, 2195-2200.

Simpson, G.B. (1984). Lexical ambiguity and its role in models of word recognition.

Psychological Bulletin, 96, 316-340.

Simpson, G.B. (1994). Context and the processing of ambiguous words. In M.A. Gernsbacher (Ed.), Handbook of Psycholinguistics (pp 359-374). San Diego: Academic Press.

Singer, M. (1994). Discourse inference processes. In M. A. Gernsbacher (Ed.), Handbook of psycholinguistics (pp. 479-516). San Diego, CA: Academic Press.

Slowiaczek, M.L., & Clifton, C. (1980). Subvocalization and reading for meaning. Journal of Verbal Learning and Verbal Behavior, 19, 573-582.

Speer, S., & Clifton, C. (1998). Plausibility and argument structure in sentence comprehension. Memory & Cognition, 26, 965-979.

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, M. J., Tanenhaus, M., Eberhard, K. M., & Sedivy, J. C. (2001). Eye movements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution. Cognitive Psychology, in press

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

Stanovich, K.E., & West, R.F. (1979). Mechanisms of sentence context effects in reading: Automatic activation and conscious attention. Memory & Cognition, 7, 77-85.

Stanovich, K.E., & West, R.F. (1983). On priming by a sentence context. Journal of Experimental Psychology: General, 112, 1-36.

Steedman, M., J., & Altmann, G. T. M. (1989). Ambiguity in context: A reply. Language and Cognitive Processes, 4, SI 77-105.

Stevenson, S. (1994). Competition and recency in a hybrid network model of syntactic dependencies. Journal of Psycholinguistic Research, 23, 295-322.

Stolz, W. (1967). A study of the ability to decode grammatically novel sentences. Journal of Verbal Learning and Verbal Behavior, 6, 867-873.

Stone, G.O., & Van Orden, G.C. (1994). Building a resonance framework for word recognition using design and systems principles. Journal of Experimental Psychology: Human Perception and Performance, 20, 1248-1268.

Stowe, L. (1986). Parsing wh-constructions: Evidence for on-line gap location. Language and Cognitive Processes, 1, 227-246.

Sturt, P., Pickering, M. J., Scheepers, C., & Crocker, M. W. (in press (2001)). The preservation of structure in language comprehension: Is reanalysis the last resort? Journal of Memory and Language.

Suh, S., & Trabasso, T. (1993). Inferences during reading: Converging evidence from discourse analysis, talk-aloud protocols, and recognition priming. Journal of Memory and Language, 32, 279-300.

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

Swinney, D.A., & Hakes, D.T. (1976). Effects of prior context upon lexical access during sentence comprehension. Journal of Verbal Learning and Verbal Behavior, 15, 681-689.

Tabossi, P. (1988). Accessing lexical ambiguity in different types of sentential contexts. Journal of Memory and Language, 27, 324-340.

Taft, M. (1991). Reading and the mental lexicon. London: Erlbaum.

Tanenhaus, M. K., Boland, J. E., Mauner, G., & Carlson, G. N. (1993). More on combinatory lexical information: Thematic structure in parsing and interpretation. In G. Altmann & R. Shillcock (Eds.), Cognitive models of speech processing: The second Sperlonga meeting (pp. 297-319). Hillsdale: Erlbaum.

Tanenhaus, M. K., & Carlson, G. N. (1990). Comprehension of deep and surface verbphrase anaphors. Language and Cognitive Processes, 5, 257-280.

Tanenhaus, M. K., Boland, J. E., Mauner, G., & Carlson, G. N. (1993). More on combinatory lexical information: Thematic structure in parsing and interpretation. In G. Altmann & R. Shillcock (Eds.), Cognitive models of speech processing: The second Sperlonga meeting (pp. 297-319). Hillsdale: Erlbaum.

Tanenhaus, M., Garnsey, S., & Boland, J. (1990). Combinatory lexical information and language comprehension. In G. Altmann (Ed.), Cognitive models of speech processing (pp. 383-408): MIT Press.

Tanenhaus, M., 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. (1996). Eye-tracking. Language and Cognitive Processes, 11, 583-588.

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

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

Tanenhaus, M. K., & Trueswell, J. C. (1995). Sentence comprehension. In J. Miller & P. Eimas (Eds.), Handbook of perception and Cognition: Speech, Language, and Communication, Second Edition (Vol. 11, pp. 217-262). San Diego: Academic Press.

Traxler, M. J., & Pickering, M. J. (1996a). 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. (1996b). Case-marking in the parsing of complement sentences: Evidence from eye movements. Quarterly Journal of Experimental Psychology, 49A, 991-1004.

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

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

Trueswell, J. C., & Tanenhaus, M. K. (1992). Consulting temporal context during sentence comprehension: Evidence from the monitoring of eye movements in reading. 14th Cognitive Science Society.

Trueswell, J. C., Tanenhaus, M. K., & Kello, C. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. Journal of Experimental Psychology: Learning, Memory, and Cognition, 19, 528-553.

Tulving, E., & Gold, C. (1963). Stimulus information and contextual information as determinants of tachistoscopic recognition of words. Journal of Experimental Psychology, 66, 319-327.

Tulving, E., Mandler, G., & Baumal, R. (1964). Interaction of two sources of information in tachistoscopic word recognition. Canadian Journal of Psychology, 18, 62-71.

Yamashita, H. (1997). The effects of word-order and case marking information on the processing of Japanese. Journal of Psycholinguistic Research, 26, 163-188.

van Dijk, T. A., & Kintsch, W. (1983). Strategies of discourse comprehension.:Academic Press.

Van Orden, G.C. (1987). A rows is a rose: Spelling, sound, and reading. Memory & Cognition, 15, 181-198.

Van Orden, G.C., & Goldinger, S.D. (1994). Interdependence of form and function in cognitive systems explains perception of printed words. Journal of Experimental Psychology: Human Perception and Performance, 20, 1269-1291.

Van Orden, G.C., Johnston, J.C., & Hale, B.L. (1988). Word identification in reading proceeds from spelling to sound to meaning. Journal of Experimental Psychology: Learning, Memory, and Cognition, 14, 371-386.

Van Orden, G.C., Pennington, B.F., & Stone, G.O. (1990). Word identification in reading and the promise of subsymbolic psycholinguistics. Psychological Review, 97, 488-522.

Van Petten, C., & Kutas, M. (1987). Ambiguous words in context: An event-related potential analysis of the time course of meaning activation. Journal of Memory and Language, 26, 188-208.

Vosse, T., & Kempen, G. (2000). Syntactic structure assembly in human parsing: A computational model based on competitive inhibition and lexicalist grammar. Cognition, 75, 105-143.

Vu, H., & Kellas, G. (1999). Contextual strength modulates the subordinate bias effect: Reply to Rayner, Binder, and Duffy. Quarterly Journal of Experimental Psychology,52A,853-856.

Warren, R.M. (1970). Perceptual restoration of missing speech sounds. Science, 167, 392-393.

Warren, R.M., & Obusek, C.J. (1971). Speech perception and phoneme restorations. Perception & Psychophysics, 9, 358-362.

Wheeler, D.D. (1970). Processes in word recognition. Cognitive Psychology, 1, 59-85.

Zwaan, R. A. (1996). Procesing narrative time shifts. Journal of Experimental Psychology: Learning, Memory and Cognition, 22, 1196-1207.

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. Psychological Science, 6, 292-297.

Acknowledgments

Preparation of this chapter was supported by a Research Scientist Award to the first author (MH01255) and by Grants HD17246, HD18708, and HD26765 from the National Institute of Health.

Table 1: Structural attachment ambiguities for which garden-pathing has been experimentally demonstrated (after Frazier & Clifton, 1996)

a. Main clause/reduced relative: *The horse raced past the barn (fell).*

b. NP vs. S complement: *John knew the answer to the physics problem was wrong /very well.*

c. direct object vs. subject of S2: *While Mary was mending the sock (it) fell off her lap.*

d. NP conjunction vs. Sent conjunction: *Jacob kissed Miriam and her sister (laughed).*

e. PP attachment to VP/NP: *Sandra wrote a letter to Mary.*

f. complement/relative clause: *John told the girl that Bill liked the story.*

g. attachment of NP as second object/relative on first object: *Fred gave the man the dog (bit the package).*

h. purpose clause vs. rationale clause: *Nixon*$_i$ *bought a 1960's version of Trivial Pursuit*$_j$ *(pro*$_j$*/pro*$_i$*) to amuse his friends.*

i. attachment of PP to lower clause/higher clause: *I put the book that you were reading in the library (into my briefcase).*

j. attachment of S to lower clause/higher clause: *Fred will realize that Mary left when the party starts/started.*

k. attachment of Adverb to lower clause/higher clause *We remembered that the assignment will be due yesterday/tomorrow.*

l. left branching vs. right branching N-N compound *The [butter cream] factory / The concrete [cream factory].*

m. (In Italian) prefer *pro* to trace reading, supporting the Minimal Chain Principle: *Ha chiamato Giovanni. ("Giovianni has called," or "Someone has called Giovanni.")*

Figure Captions

Figure 1. Examples of eye-contingent display change paradigms. The top line shows a line of normal text. In the Moving Window example, a window of readable text (in the example extending 9 letter spaces to the left and to the right of the fixation point) is available on each fixation (with letters outside the window replaced by x's). Two consecutive fixations are shown. In the Boundary example, the word tune is initially present in the text, but when the reader's eye movement crosses an invisible boundary location, tune changes to song. In the Fast Priming example, a nonword letter string (rcsp) initially occupies the target word location, but when the reader's eye movement crosses the boundary location, the prime word tune is presented for about 30-35 ms and then replaced by the target word song (which is visible for the rest of the trial).

Figure 2. A Dual-route model of word recognition of Coltheart, Curtis, Atkins, and Haller (1993). Copyright 1993 by the American Psychological Association. Reproduced by permission.

Figure 3. A Connectionist model of word recognition of Seidenberg and McClelland (1989). Copyright 1989 by the American Psychological Association. Reproduced by permission.

Figure 4. A sample display in the word superiority effect paradigm.

Figure 5. Successive structures created by application of phrase structure rules.

Normal Text:

```
John composed a new song for the children
```

Moving Window:

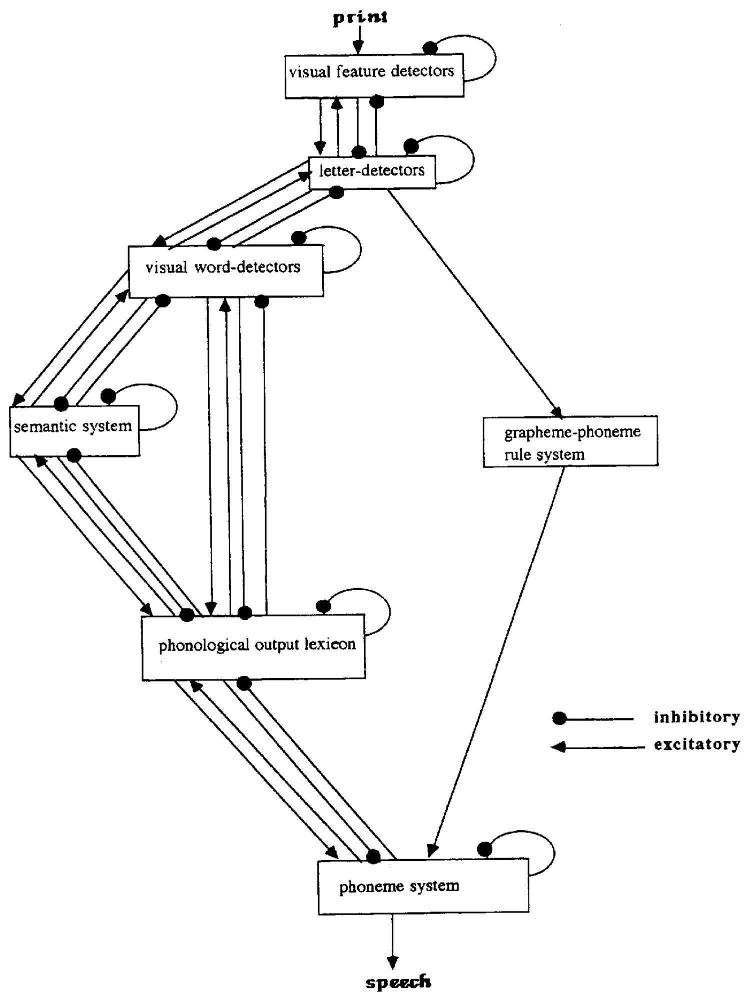
```
xxxx xxxxxxxd a new song for thx xxxxxxx
                     *
xxxx xxxxxxxx x xew song for the chixxxxx
                     *
```

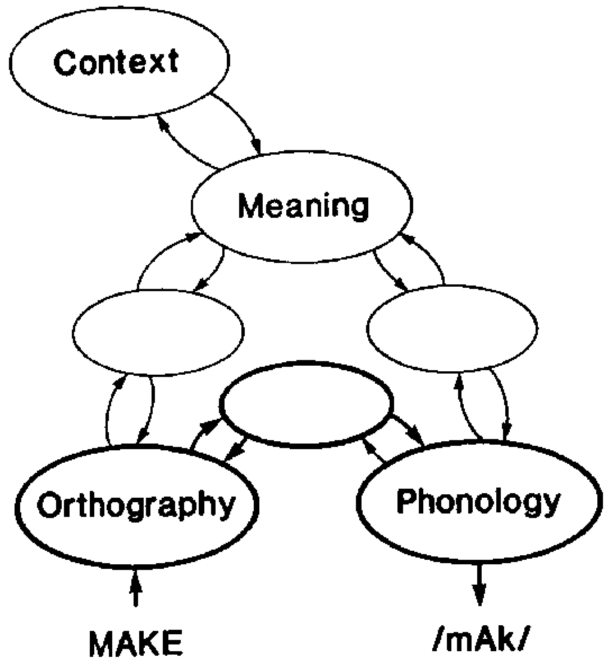
Boundary:

```
John composed a new tune for the children
                     *
John composed a new song for the children
                     *
```

Fast Priming:

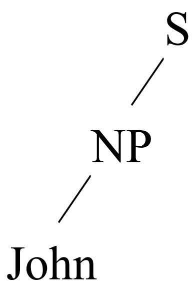
```
John composed a new rcsp for the children
                     *
John composed a new tune for the children
                     *
John composed a new song for the children
                     *
```


print
visual feature detectors
letter-detectors
visual word-detectors
semantic system
grapheme-phoneme rule system
phonological output lexicon
inhibitory
excitatory
phoneme system
speech

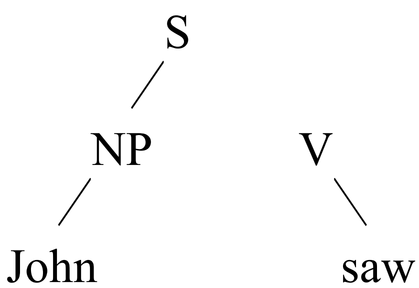

Context
Meaning
Orthography
Phonology
MAKE
/mAk/

•	Fixation point
word	Stimulus display
d #### k	Response choices

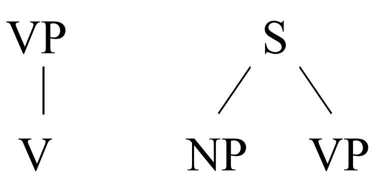
1.



2.



3. Access templates:



4.