Eye Movements and Lexical Access in Spoken-Language Comprehension: Evaluating a Linking Hypothesis between Fixations and Linguistic Processing

Michael K. Tanenhaus,[1,3] James S. Magnuson,[1] Delphine Dahan,[2] and Craig Chambers[1]

A growing number of researchers in the sentence processing community are using eye movements to address issues in spoken language comprehension. Experiments using this paradigm have shown that visually presented referential information, including properties of referents relevant to specific actions, influences even the earliest moments of syntactic processing. Methodological concerns about task-specific strategies and the linking hypothesis between eye movements and linguistic processing are identified and discussed. These concerns are addressed in a review of recent studies of spoken word recognition which introduce and evaluate a detailed linking hypothesis between eye movements and lexical access. The results provide evidence about the time course of lexical activation that resolves some important theoretical issues in spoken-word recognition. They also demonstrate that fixations are sensitive to properties of the normal language-processing system that cannot be attributed to task-specific strategies.

KEY WORDS:

EYE MOVEMENTS AND SPOKEN-LANGUAGE COMPREHENSION

A rapidly expanding community of psycholinguists is now using head-mounted eye trackers to study spoken-language comprehension, and, more recently, language production (e.g., Eberhard, 1998; Griffin & Bock, 2000, Meyer *et al.,* 1998). The first presentation of research using head-mounted eye tracking

---

was at the 1994 CUNY meeting. At the 2000 meeting, five talks and seven posters from four different laboratories presented research using this technique. The topics ranged from how temporarily misleading coarticulatory cues in vowels affect lexical access (Tanenhaus, Dahan, Magnuson, & Hogan, 2000) to how, and if, children take into account information about common ground in reference resolution (Hurewitz, Brown-Schmidt, Gleitman, & Trueswell, 2000; Nadig & Sedivy, 2000). Other topics focused on more classical questions in sentence processing, including what referential domains listeners consider when interpreting reflexives and pronouns (Runner, Sussman, & Tanenhaus, 2000); how argument structure is used in interpreting sentences with filler-gap dependencies (Sussman & Sedivy, 2000), and whether listeners use lexical conceptual knowledge to make predictions about upcoming phrases (Altmann, Haywood, & Kamide, 2000; Kako & Trueswell, 2000).

In the typical "visual world" study of comprehension, the participant follows instructions to look at, pick up, or move, one of a small set of objects presented in a well-defined visual workspace (Tanenhaus & Spivey-Knowlton, 1996). The timing and pattern of fixations to potential referents in the visual display is used to draw inferences about comprehension. The use of eye movements in spoken-language comprehension was pioneered in a remarkable article by Cooper (1974), who demonstrated that participants' eye movements to pictures were closely time locked to relevant information in a spoken story. The recent surge of interest in head-mounted eye tracking in psycholinguistics began with a short report by Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy (1995) who examined syntactic ambiguity resolution using a task in which participants followed spoken instructions to manipulate objects in a visual workspace.

Interest in the head-mounted eye-movement paradigm has been growing for several reasons. First, eye movements provide a continuous measure of spoken-language processing in which the response is closely time locked to the input without interrupting the speech stream. Eye movements provide insights into the time course of reference resolution (Eberhard, Spivey-Knowlton, Sedivy, & Tanenhaus 1995; Sedivy, Tanenhaus, Chambers, & Carlson, 1999; Trueswell, Sekerina, Hill, & Logrip, 1999; Altmann & Kamide, 1999; Arnold, Eisenband, Brown-Schmidt & Trueswell, 2000), while providing sufficient temporal resolution to measure lexical access in continuous speech (Allopenna, Magnuson, & Tanenhaus, 1998; Spivey & Marian, 1999; Spivey-Knowlton, 1996; Dahan, Magnuson, Tanenhaus & Hogan, in press). Second, the eye-movement paradigm can be used with natural tasks that do not require metalinguistic judgments. Thus, it is well-suited for studies with young children (Trueswell *et al.,* 1999) and with brain-damaged populations (Yee, Blumstein, & Sedivy, 2000). Third, the presence of a visual world makes it possible to ask questions about real-time interpretation, especially questions about semantic

reference, that would be difficult to address, and perhaps intractable, if one were limited to measures of processing complexity for written sentences or spoken utterances (cf. Sedivy *et al.,* 1999). Finally, and perhaps most importantly, the paradigm allows one to study real-time language production and comprehension in natural tasks involving conversational interaction. This makes it possible to bridge the two dominant traditions in language-processing research: the "language-as-action" tradition, which has focused on natural interactive conversation while generally ignoring questions about the mechanisms underlying real-time language processing and the "language-as-product" tradition, which has focused on the time course of processing while being primarily limited to "decontextualized language" (Clark, 1992).

The remainder of this article is structured as follows. In the second section, we briefly review results demonstrating that referential information, including properties of referents relevant to specific actions, influences even the earliest moments of syntactic processing. In the next section, we identify three concerns about the head-mounted eye-tracking paradigm that we believe are shared by many in the sentence-processing community. These are concerns about (1) task-specific strategies, which may arise because the studies use a circumscribed visual context, and (2) the linking hypotheses between eye movements and linguistic processing. In the fourth section, we address these concerns by reviewing a recent line of research in which we have been using eye movements to investigate the time course of lexical access in continuous speech. This work has both a theoretical and a methodological focus. On the theoretical side, we have been examining how fine-grained phonetic information affects the time course of activation of lexical candidates during lexical access. On the methodological side, we have been developing and evaluating a specific linking hypothesis between underlying lexical activation and eye movements. By focusing directly on the linking hypothesis, we also address concerns about the use of a circumscribed world and about task sensitivity.

VISUAL CONTEXT AND SYNTACTIC AMBIGUITY RESOLUTION

One of the first applications of head-mounted eye-tracking to sentence processing examined syntactic processing of spoken sentences containing temporarily ambiguous phrases in visually defined referential contexts (Tanenhaus *et al.,* 1995; Spivey, Tanenhaus Eberhard, & Sedivy, in press). Crain and Steedman (1985) called attention to the fact that many of the classic structural ambiguities involve a choice between a syntactic structure, in which the ambiguous phrase modifies a definite noun phrase, and one in which it is a syntactic complement or argument of a verb phrase. Under these

conditions, the complement analysis is typically preferred. For instance, in example (1), readers and listeners will initially misinterpret the prepositional phrase, "on the towel," as the goal argument of "put" rather than as an adjunct modifying the noun phrase, "the apple," resulting in a garden path.

    1. Put the apple on the towel in the box.

    Crain and Steedman noted that one use of modification is to differentiate an intended referent from other alternatives. For example, it would be odd for example (1) to be uttered in a context in which there was only one perceptually salient apple, such as the scene in Figure 1 (Panel A), whereas it would be natural in contexts with more than one apple, as in the scenes illustrated in Panels B and C. In these contexts, the modifying phrase, "on the towel," provides information about which of the apples is intended. Crain and Steedman proposed that listeners might initially prefer the modification analysis to the complement analysis in situations that provided the appropriate referential context. Moreover, they suggested that referential fit to the context, rather than syntactic complexity, was the primary factor controlling syntactic preferences.

    Numerous empirical studies have now been conducted to evaluate the extent to which initial parsing decisions are influenced by referential context, beginning with studies by Altmann and Steedman (1988) and Ferreira and Clifton (1986). (For more recent reviews, see Altmann, 1998; Gibson
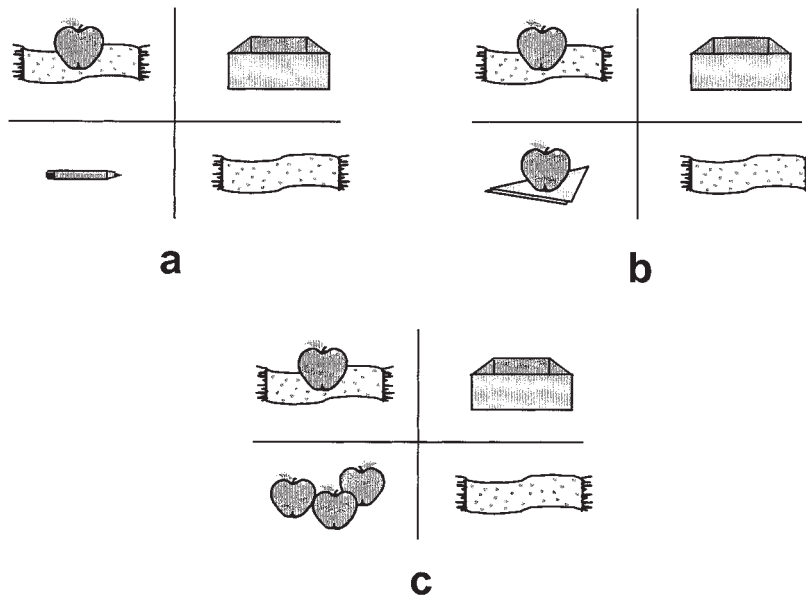

Fig. 1. Sample displays from Spivey *et al.* (in press).

& Pearlmutter, 1998; Spivey & Tanenhaus, 1998; Tanenhaus & Trueswell, 1995.) Nearly all of these studies have used text typically presenting stimuli in a short paragraph. A referential context is created by setting up a scenario, which introduces one or more potential referents for a definite noun phrase in a subsequent target sentence, which contains a temporarily ambiguous phrase. Reading time for critical regions of the target sentence are used to infer whether the context influenced how the reader first parsed the ambiguous phrase. Text has been used because the theoretical questions required response measures that can provide fine-grained temporal information about ambiguity resolution. Self-paced reading, and especially eye tracking during reading, provide the necessary grain because processing difficulty can be measured for each word in a sentence (Rayner, 1998).

Studies of syntactic ambiguity resolution using reading paradigms have provided, and continue to provide, valuable information about the role of context in sentence processing. However, they also have some intrinsic limitations. One limitation is that reading-time measures provide only a general measure of processing difficulty. That is, they do not provide information about *what* is being processed, or *how* it is being processed, but merely indicate whether the processing requires additional time compared to some baseline. A second limitation is that context can only be created by evoking events and entities through linguistic expressions, which must be held in memory. However, it is widely known that the relevant notion of "context" for an utterance includes not only previous discourse but also the entities in the interlocutors' environment, as well as the set of presuppositions shared by discourse participants, including those created by the unfolding utterance (cf. Clark, 1992). It is important not to conflate the more general questions of how, and when, a context can influence initial syntactic processing with the narrower question of how linguistically introduced context influences syntactic processing of a subsequent sentence.

Spivey *et al.* (in press) investigated the processing of temporarily ambiguous sentences such as (1), repeated as (2a), and unambiguous control sentences, such as (2b), in contexts such as the ones illustrated in Figure 1. The objects illustrated in the figures were placed on a table in front of the participant. Participants' eye movements were monitored as they performed the action in the spoken instruction.

2. a. Put the apple on the towel in the box.
   b. Put the apple that's on the towel in the box.

The results provided striking evidence for immediate use of the visual context. In the one-referent context (Panel A), participants looked at the false goal (the empty towel) on fewer than 10% of the trials with the unambiguous instructions. In contrast, participants looked at the false goal on more than 50% of the trials with the temporarily ambiguous instruction. Detailed analysis of

the timing of these fixations indicated that they began as the word "towel" was uttered. These results provide clear evidence that participants were garden pathed with the ambiguous instruction and momentarily shifted their attention to the false goal. This result is consistent with the argument preference predicted by all structurally based models of ambiguity resolution.

However, looks to the false goal were dramatically reduced in the two-referent context (Panel B). Crucially, there was not even a suggestion of a difference between the proportion of looks to the false goal with the ambiguous and the unambiguous instructions. Moreover, the timing of the fixations provided clear evidence that the prepositional phrase was being immediately interpreted as modifying the NP. Participants typically looked at one of the potential referents as they heard the beginning of the instruction, e.g., "put the apple." On trials in which participants looked first at the incorrect Theme (e.g., the apple on the napkin), they immediately shifted to the correct Theme (the apple on the towel) as they heard "towel." The timing was identical for the ambiguous and unambiguous instructions (see Trueswell *et al.,* 1999 for similar results).

The condition illustrated in Panel C provides important additional information. In this condition, modification is felicitous. In fact it would be markedly odd to ask someone to, "Put the apple in the box," rather than, "Put the apple (that's) on the towel in the box." Nonetheless, use of a definite noun phrase (e.g., "put the apple . . .") strongly biases the listener toward the single referent (the apple on the towel) rather than toward the group (the three apples). In this condition, participants initially fixated on the single referent, but showed no tendency to look at the false goal.

Recently, Chambers, Tanenhaus, & Magnuson (2000) have shown that real-world knowledge that is relevant to specific actions also modulates attachment preferences. Chambers *et al.* used temporarily ambiguous instructions such as, "Pour the egg in the bowl over the flour," and unambiguous instructions, such as, "Pour the egg that's in the bowl over the flour," with displays such as the one illustrated in Figure 2.

The critical manipulation was whether one or both potential referents (e.g., the two eggs) matched the affordances required by the action denoted by the verb. In the example, one can *pour* a liquid egg, but not a solid egg. When both potential referents matched the verb (e.g., the condition with two liquid eggs, as in Panel A), there were few looks to the false goal (e.g., the bowl) and no differences between the ambiguous and unambiguous instructions. Thus, the prepositional phrase was correctly interpreted as a modifier, replicating the pattern observed by Spivey *et al.* (in press; also see Tanenhaus *et al.,* 1995, 1999). However, when the properties of only one of the potential referents matched the verb, (e.g., the condition where there was a liquid egg and a solid egg, as in Panel B), participants were far more likely to look to the false goal (the bowl) with the ambiguous instruction than with the unambiguous instruction. Listeners were now garden pathed by the ambiguous instruction because
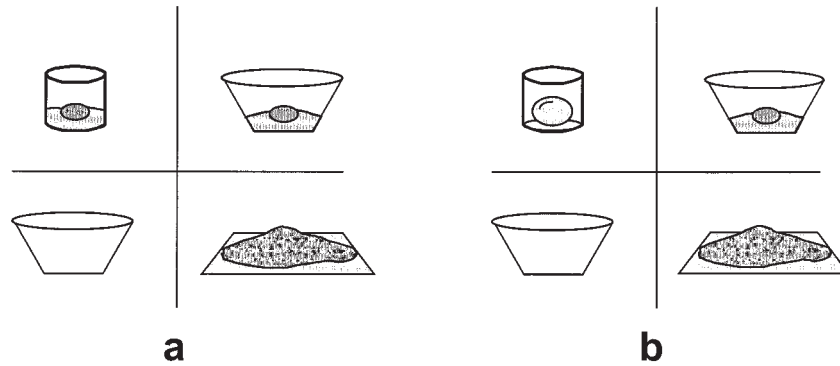
Fig. 2. Sample displays from Chambers *et al.* (2000).

there was only one pragmatically appropriate referent (the liquid egg) compatible with a pouring action, showing the same pattern of fixations as Spivey *et al.* found in their one-referent condition (Figure 1, Panel A).

These results have important implications for our models of real-time sentence processing and for how we study language processing. The results show that visually co-present contexts modulate initial attachment preferences, reversing attachment preferences even when the otherwise preferred attachment is an obligatory argument and the less-preferred attachment is an optional adjunct. Any serial parsing models that have been proposed to date cannot accommodate these results. Moreover, the relevant referential domain for syntactic processing immediately takes into account context-specific real-world knowledge[4] making the results difficult for modular theories to accommodate. Given these results, approaches to language comprehension that assign a central role to encapsulated linguistic subsystems are unlikely to prove fruitful. More promising are theories in which grammatical constraints are integrated into processing systems that continuously coordinate linguistic and nonlinguistic information as the linguistic input is processed.

However, we know that many in the sentence-processing community are resistant to these conclusions because of serious methodological questions about the eye-tracking paradigm. In the next section, we discuss two of these concerns.

---

[4] One could argue that real world knowledge came into play only because the lexical conceptual properties of the verb "pour" placed constraints on the properties of its theme argument. The strongest test of the claim that real-world knowledge is consulted requires use a verb that does not plausibly contain any such constraints (e.g., "put") and that can be defined independently of the specific context. For example, a participant holding a pair of tongs could only use the tongs to move a solid egg but not a liquid egg, in following an instruction, such as "Put the egg in the bowl on the flour." Research examining conditions like these is currently in progress.

TWO CONCERNS ABOUT HEAD-MOUNTED EYE TRACKING

Task-Specific Strategies

Perhaps the most serious concern is that the combination of a circumscribed visual world and restricted set of instructions encourages participants to develop task-specific strategies that bypass "normal" language processing. The argument goes as follows. Upon viewing a scene, such as the one illustrated in Figure 1 (Panel B), the participant might note that there are two apples and predict that the upcoming instruction is likely to focus on one of the apples, and perhaps even predict the form of the instruction. However, this type of prediction is unlikely in most real-world environments because the immediate environment is rarely this circumscribed, even in task-oriented dialog. Moreover, people are not limited to talking about the immediate environment. Thus, at best, the results described in second section (Visual Context and Syntactic Ambiguity Resolution) may not scale up to more realistic environments. At worst, they may be due to a specific problem-solving strategy induced by the task. While it is difficult to rule out this kind of argument, in principle, we are skeptical about its validity for several reasons. First, in all of our experiments we construct the pairing of scenes and instructions to avoid predictable contingencies. Second, the results do not seem to depend upon lengthy exposure to the visual context. The same pattern of results occurs when participants have twenty seconds or so to view the scene prior to the beginning of the instruction and when the scene is presented less than a second before the instruction begins (e.g., Sedivy *et al.,* 1999). Third, participants deny consciously encoding the scene, naming the objects, or anticipating the instructions. If participants were strategically engaging in any of these behaviors, we might expect some level of awareness. Fourth, participants' claims that they are not encoding the scene linguistically are consistent with the growing literature on scene perception and "change blindness." This literature shows that people do not consciously encode and represent details about even relatively simple scenes (cf., Simons, 2000). Finally, and most convincingly, eye movements are affected by properties of the linguistic system that would not come into play if strategies were allowing listeners to bypass the general language-processing system (see sections on Time Course of Frequency Effects and Subcategorical Mismatches).

The Linking Hypothesis

The interpretation of all behavioral measures depends upon a theory, or "linking hypothesis," that maps the response measure onto the theoretical

constructs of interest. For example, using cross-modal semantic priming to infer that a noun phrase has been linked to an empty category (e.g., Nicol & Swinney, 1989), depends upon the following chain of inference. Encountering an anaphor triggers a memory search that identifies the antecedent. The antecedent is then "reactivated." As a result, activation spreads to related or associated lexical concepts, facilitating their recognition. Linking hypotheses are often stated informally or are left implicit. Nonetheless, they are a necessary part of the inference chain that links theory to data; there are no "signature" patterns that provide a direct window into underlying cognitive processes. As researchers begin to address increasingly fine-grained questions about cognitive microstructure, explicit and quantitative linking hypotheses take on added importance. Tanenhaus, Spivey-Knowlton, and Hanna (2000) developed this argument in detail focusing on context effects on reading times for locally ambiguous sentences.

When the duration and pattern of fixations are used to study sentence processing in reading, the linking hypothesis between fixation duration and underlying processes is intuitive; reading times increase when processing becomes more difficult. Nonetheless, our theories of sentence processing will eventually have to combine explicit models of the underlying processes with explicit models of how these processes affect fixations. Progress will depend upon developing and refining these models rather than on our intuitions about tasks. The link between fixation patterns and spoken-language comprehension is less intuitive than the link between fixations and comprehension in reading. Informally, we have automated behavioral routines that link a name to its referent; when the referent is visually present and task relevant, then recognizing its name accesses these routines, triggering a saccadic eye movement to fixate the relevant information. As in reading, we do not yet have models of sentence processing and models of eye movements that are explicit enough to generate fixations with enough accuracy to capture much of the variance. Thus we must rely on hypothesis-testing experiments and underspecified linking hypotheses. However, in more restricted domains, such as word recognition, where models are more explicit and there are fewer degrees of freedom, it is possible to develop and test more formal versions of our linking hypotheses. To the extent that these efforts are successful, it increases our confidence in the response measure and its linking hypothesis.

We now turn to a review of some of our recent work using eye movements to examine the time course of lexical access in continuous speech. This work was primarily motivated by theoretical questions in spoken-word recognition. However, a secondary motivation was that spoken-word recognition is a natural domain for evaluating concerns about task-specific strategies and linking hypotheses in the visual world paradigm. Current models make explicit quantitative predictions about the time course of

activation, which allow us to develop and test an explicit linking hypothesis. Moreover, clear evidence that eye movements can be used to trace the time course of lexical activation within words in continuous speech, including effects of fine-grained phonetic information, and subtle effects of lexical competitors that are difficult to capture with other response measures makes it highly implausible that the eye-movement paradigm would be too coarsely grained to detect effects of temporary syntactic misanalysis.

EYE MOVEMENTS AND LEXICAL ACCESS IN CONTINUOUS SPEECH

As the sound pattern of a spoken word unfolds over time, recognition takes place against a backdrop of partially activated alternatives that compete for recognition. The most activated alternatives are those that most closely match the input. For instance, as a listener hears the word "candy," lexical representations of words with similar sounds, such as "candle," will also be activated. The number of competitors, their frequency of occurrence in the language, as well as the frequency of occurrence of the target word itself, all affect recognition (e.g., Luce & Pisoni, 1998; Marslen-Wilson, 1987, 1990). Many of the central theoretical questions about spoken word recognition focus on details about the time course of activation among lexical competitors. As researchers begin to address these questions, it becomes increasingly important to have response measures that are sensitive to time course. In addition, evaluating competing theoretical proposals requires explicit models and explicit quantitative linking hypotheses that map predictions from models onto behavioral response measures.

In some of our earliest visual world studies (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1996), we demonstrated that the mean latency to fixate a referent was delayed when the visual workspace contained a "cohort" competitor whose name shared the same initial onset and vowel (e.g., when the workspace contained both a candle and a candy). Moreover, when the first fixation was to an object other than the referent, participants were far more likely to look at the competitor than an unrelated object. These results encouraged us to explore the feasibility of using the eye movement paradigm as a tool for examining lexical access in continuous speech. We have been pursuing a strategy of addressing unresolved theoretical issues about the time course of lexical access in continuous speech while, at the same time, developing and evaluating a linking hypothesis between lexical activation and eye movements.

Cohort and Rhyme Competition and An Explicit Linking Hypothesis

Allopenna *et al.* (1998) evaluated the time course of activation for lexical competitors that shared initial phonemes with the target word (e.g., beaker and beetle) or that rhymed with the target word (e.g., beaker and speaker). The Cohort Model and its descendents (e.g., Marslen-Wilson, 1987, 1990, 1993) assume that any featural mismatch is sufficient to strongly inhibit a lexical candidate. On this view, as "beaker" unfolds, the lexical representations of "beaker" and "beetle" become activated, but not the lexical representation of "speaker." In contrast, continuous mapping models, such as TRACE (McClelland & Elman, 1986) predict that both cohort and rhyme competitors will become active.

In the Allopenna *et al.* studies, participants were instructed to fixate a central cross and then followed a spoken instruction to move one of four objects displayed on a computer screen with the computer mouse (e.g., "Look at the cross. Pick up the beaker. Now put it above the square"). A sample display is presented in Figure 3, Panel A. TRACE simulations of the activation of the lexical representations of the pictured objects for spoken word "beaker" are presented in Figure 3, Panel B.

Eye movements to each of the objects were recorded as the name of the referent object unfolded over time. The probability of fixating each object as the target word unfolded was hypothesized to be closely linked to the activation of the lexical representation of this object (i.e., its name). The assumption providing the link between lexical activation and eye movements is that the activation of the name of a picture determines the probability that a subject will shift attention to that picture and thus make a saccadic eye movement to fixate it.[5]

Allopenna *et al.* (1998) formalized this linking hypothesis by using the Luce (1959) choice rule to convert activations at each moment in processing in TRACE to predictions about the proportion of fixations to each of the displayed alternatives. The activation of each alternative is converted into a response strength using the following equation:

$$S = e^{ka_i}$$

S is the response strength for each item, *a* is the activation taken from TRACE, and *k* is a free parameter that determines the amount of separation

---

[5] By using the word "attention," we do not intend to suggest that participant's are consciously shifting attention. In fact, people are typically unaware of making eye movements and even of exactly where they are fixating. One possibility is that the attentional shifts take place at the level of unconscious visual routines that support accessing information from a visual scene (Hayhoe, 2000).
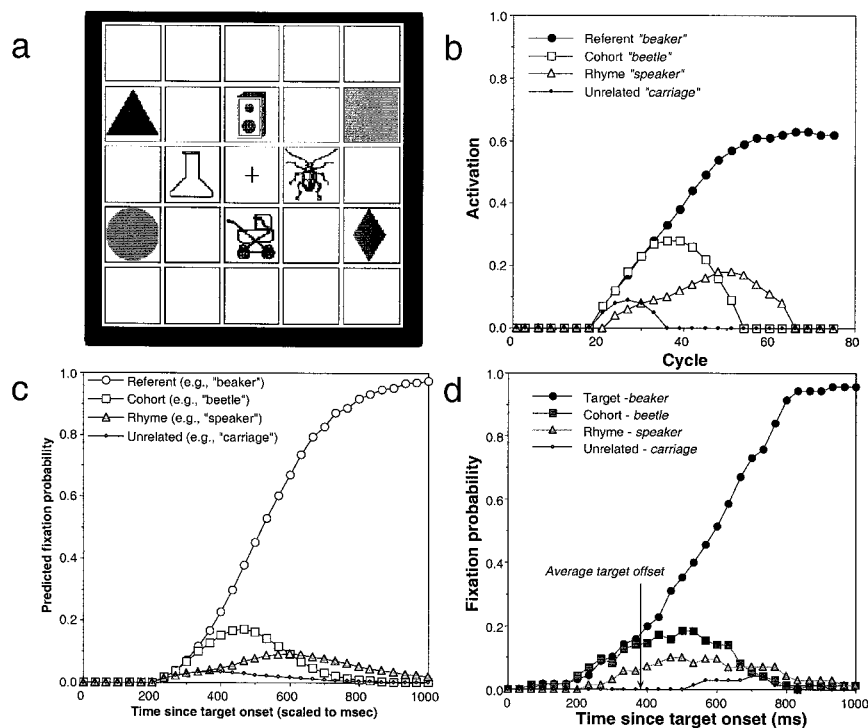
Fig. 3. Materials, simulations and results from Allopenna *et al.* (1998). Panel a, sample stimulus display from a critical trial. Panel b, TRACE activations for the four items of interest. Panel c, TRACE activations converted to predicted fixation probabilities (see text for details). Panel d, observed fixation probabilities for each item of interest.

between activation levels. The Luce choice rule is then used to convert the response strengths to response probabilities, using the following equation:

$$L_i = \frac{S_i}{\sum S_j}$$

In adopting this equation we are assuming that the activations that determine response strength come from the lexicon as a whole, but response selection is based only on the displayed alternatives.

The Luce choice rule assumes that a choice will be made at each choice point. Thus, each alternative is equally probable when there is no information. When the initial instruction is "look at the cross" or look at picture X, we scale the response probabilities to be proportional to the

amount of activation at each time step using the following equations, where max_t is the maximum activation at a particular time step, m is a constant equal to the maximum expected activation (e.g., 1.0), i is a particular item, and d_t is the scaling factor for time step t:

$$d_t = \frac{\max_t}{m}$$

$$R_i = d_t L_i$$

Thus the predicted fixation probability is determined both by the amount of evidence for an alternative and the amount of evidence for that alternative compared to the other possible alternatives. Finally, we introduce a 200-ms delay because programming an eye movement takes approximately 200 ms (Matin, Shao, & Boff, 1993). When the linking hypothesis is applied to TRACE simulations of activations for the stimuli used by Allopenna *et al.,* it generates the predicted fixations over time, shown in Figure 3 (Panel C). The data are presented in Figure 3, Panel D.

In the Allopenna *et al.* study, the proportion of fixations to referents and cohort competitors began to increase 200 ms after word onset. Thus, the eye movements were sensitive to changes in lexical activation from the onset of the spoken word. Eye movements generated early in the target word were equally likely to result in fixations to the cohort competitor (e.g., "beetle") and to the referent (e.g., "beaker"), and were more likely to result in fixations to these pictures than to distractor controls that were phonologically unrelated to the target word (e.g., "carriage"). The fixations over time to the target, the cohort competitor, and a rhyme competitor (e.g., "speaker") closely matched the predictions generated by our simple hypothesis linking activation levels in TRACE to fixation probabilities over time.

Although the Allopenna *et al.* (1998) results provide strong support for the linking hypothesis, they do not provide direct evidence that the eye-movement paradigm is sensitive to properties of the general lexicon. Moreover, one could argue that some aspects of the design and procedure might have encouraged task-specific strategies. For example, participants were told the name of each picture before the experiment began, pictures were repeated several times throughout the experiment (although distractor trials were carefully controlled to prevent probability matching), and the display was presented for about 2 before the spoken instruction began, during which time participants could scan the display. Thus one might argue that participants began each trial by encoding the scene and holding it in working memory, or even by implicitly naming the pictures. The name of the referent in the input would then be matched against these generated

representations, somehow bypassing normal lexical processing. Participants inform us that they do not name the pictures nor consciously try to encode the display. Nonetheless, it was important to provide direct empirical evidence that fixations are influenced by characteristics of the lexicon that are not directly represented in the set of pictures displayed on a trial. Such evidence is presented in the next two sections.

Time Course of Frequency Effects

Dahan, Magnuson, and Tanenhaus (in press, a) conducted two experiments using eye movements to examine the time course of frequency effects on lexical access. Previous research has shown that well-documented frequency and neighborhood effects in word recognition can be dramatically reduced or even disappear in closed-set tests (Pollack, Rubenstein, & Decker, 1959; Sommers, Kirk, & Pisoni, 1997). Thus, demonstration of a frequency effect would provide strong support for our assumption that lexical activations are not restricted to the set of displayed alternatives.

In Experiment 1, participants were presented with displays consisting of a referent along with two cohort competitors that varied in frequency and an unrelated distractor. For example, in the sample display, illustrated in Figure 4 (Panel A), the referent "bench" was presented along with the high-frequency cohort, "bed," the low-frequency cohort, "bell," and the distractor, "lobster." The procedure differed from the one used in Allopenna *et al.* in three important ways. Participants had no exposure to the pictures prior to the experiment, no pictures or names were repeated, and subjects had only 500-ms exposure to the display before the instruction began. Participants were instructed to pick up the designated object by clicking on it with the computer mouse (e.g., "Pick up the bench"). As the target word unfolded, the cohorts were expected to be fixated more than the distractor, as a consequence of their phonological similarity with the initial portion of the input. In addition, if fixations reflect lexical processing, more fixations to the high-frequency cohort than to the low-frequency cohort would be expected. Crucially, if lexical frequency operates on the lexical-access process (rather than as a possible response bias after recognition is completed), the advantage for fixating the high-frequency over the low-frequency competitor should be observed before the auditory input provides disambiguating information.

Figure 4 (Panel B) shows the proportion of fixations to the high- and low-frequency cohorts. When the picture was displayed, participants typically made a saccade to fixate on one of the pictures. Thus, at the onset of the referent in the instruction, participants were equally likely to be fixating one of the four pictures. Over the 200 to 500-ms window, the high-frequency cohort was fixated more than the low-frequency one. The fixations to high-
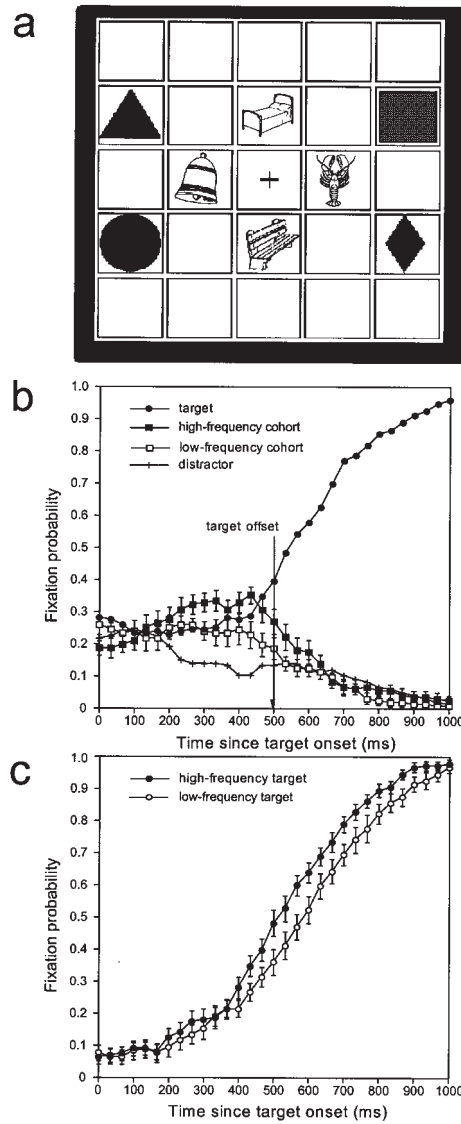
Fig. 4: Materials and results from Dahan *et al.* (in press, a). Panel a, sample stimulus display from a critical trial, with a low-frequency target (*bench*), low-frequency cohort (*bell*), and a high-frequency cohort (*bed*). Panel b, fixation probabilities over time for the high- and low-frequency cohort condition. Panel c, fixation probabilities for high- and low-frequency targets presented among unrelated distractors.

and low-frequency cohorts started diverging at about 267 ms. The fixations to the target and to the low-frequency cohort remained comparable until 467 ms after target onset (this result was expected because the frequencies of the target word and the low-frequency cohort were similar); after this point, the fixations to the target surpassed all other fixations. Measurements of the stimuli indicated an overlap of roughly 220 ms before the coarticulatory information began (e.g., the duration of the vowel before nasal information in "bench" was about 195 ms long). If one takes into account the delay for launching an eye movement (approximately 200 ms), the time locking of fixations to targets and competitors is quite remarkable.

In Experiment 2, we varied the frequency of the referent and presented it along with three phonologically unrelated distractors. We also introduced a change in procedure to minimize the proportion of trials on which the participant would be fixating on the target as when the instruction began. The instruction was composed of two parts. First, participants were asked to point to one of the distractor pictures using the computer mouse (e.g., "Point to the sock"). After a delay of 300 ms, allowing participants to move the mouse cursor to the distractor picture, they were instructed to point to the target picture (e.g., "now the bed"). Then, they were asked to move the target picture above or below one of the geometrical shapes (e.g., "Click on it and put it above the circle"). Once this was accomplished, the next trial began. If the probability of fixating a picture reflects the activation of the lexical representation associated with this picture, fixations should reach the referent pictures with high-frequency names faster than the referent pictures with low-frequency names. The results, presented in Figure 4 (Panel C), confirmed this prediction.

Pictures corresponding to high-frequency items were fixated more quickly than those corresponding to low-frequency items (563 vs. 625 ms). From about 400 ms after target onset, the proportion of fixations to the high-frequency target surpassed the proportion of fixations to the low-frequency targets, indicating that participants fixated the high-frequency earlier than the low-frequency target. Again, this result reveals the influence of frequency early in the word-recognition process.

Dahan *et al.* (in press, a) also showed that simulations with TRACE using the same parameters as those used by Allopenna *et al.* (1998) provided close fits to the data. Moreover, these simulations provided insights into the locus of frequency effects. Most crucially, however, for our current purposes, the results provide strong support for our linking hypothesis and make it highly unlikely that subjects are adopting a special verification strategy that bypasses normal lexical processing. If the lexical candidates that entered the recognition process were restricted to the visually present alternatives, then we would not expect to see effects of frequency. This is

especially true for Experiment 2, where we found dear frequency effects even when the display did not contain competitors with names that were similar to the referent. However, truly compelling evidence that we are observing effects of lexical activation from the general lexicon requires demonstrating that fixations to a referent are influenced by lexical competitors that are neither named nor pictured. The next section presents just such a demonstration.

Subcategorical Mismatches: Effects of Nondisplayed Alternatives

Dahan, Magnuson, Tanenhaus, and Hogan (in press) examined the time course of lexical competition when mismatching coarticulatory cues (i.e., inconsistent with the actual identity of the following consonant) match a word in the lexicon compared to when these cues do not match an existing word. All current models of spoken-word recognition predict that the coarticulatory cues matching an existing word will temporarily favor this word, thus disfavoring the word that will eventually match the entire sequence (i.e., the target word). Marslen-Wilson and Warren (1994) presented evidence that they argued was inconsistent with competition operating via lateral inhibition, as in models like TRACE. They created cross-spliced word sequences whose initial CV portion had been excised from another token of the same word [e.g., *jo*(b) + (jo)*b,* W1W1 condition], from another existing word [e.g., *jo*(g) + (jo)*b,* W2W1 condition], or from a nonword [e.g., *jo*(d) + (jo)*b,* N3W1 condition]. For the W2W1 [jo(g)b] and N3W1 [jo(d)b] conditions, formant transitions in the vowel provide misleading information about the place of articulation of the following consonant. Thus, these stimuli contained *subcategorical* phonetic mismatches (Streeter & Nigro, 1979; Whalen, 1984, 1991).

Marslen-Wilson and Warren reasoned that, if lexical candidates inhibit one another, as predicted by TRACE, then lexical decisions to words with subcategorical mismatches cross-spliced from words should be slower than lexical decisions to the words cross-spliced from nonwords. In TRACE, and similar models such as Shortlist (Norris, 1994), for W2W1, the initially activated competitor W2 (e.g., *jog*) inhibits the target W1 (e.g., *job*); in N3W1, this inhibition is substantially reduced because coarticulatory cues in the vowel from the nonword N3 (e.g., *jod*) only weakly support both W2 and W1. Inhibition modifies the activations of words throughout processing. Thus, the degree to which the competitor (W2) is activated affects the activation of the target (W1) throughout the recognition process. However, they found no effect of the lexical status of the conflicting cues on lexical decision latencies. Responses to W2W1 cross-spliced sequences [e.g., jo(g)b], containing coarticulatory information in the vowel coming from a lexical competitor (e.g., *jog*), were equivalent to responses to N3W1 cross-spliced

sequences [e.g., jo(d)b], containing coarticulatory information coming from a nonword *jod* (N3). Marslen-Wilson and Warren (1994) interpreted this result as evidence against inhibition between activated word units, as instantiated in the TRACE and Shortlist models.

More recently, Norris *et al.* (in press; see also McQueen, Norris, & Cutler, 1999) showed that the absence of difference between W2W1 and N3W1 can fall out of a model incorporating lateral inhibition if the lexical competition between W2 and W1 is resolved before lexical-decision responses are generated. These simulations illustrate how difficult it is to distinguish among competing models without detailed information about the time course of activation of lexical competitors. The different patterns of activation predicted by models with and without active inhibition might occur too early in processing to be detected using lexical decisions.

Dahan *et al.* (in press, b) monitored participants' eye movements to pictured objects as they heard the referent's name in each of three splicing conditions, W1W1, W2W1, and N3W1. In order to 'minimize the proportion of trials where participants were already fixating the target picture at the onset of the target word, participants were first instructed to point with the mouse cursor to one of the distractor pictures (e.g., "Point to the bass"). As soon as the cursor reached the picture, the critical instruction containing the target word was played (e.g., "now the net"). The test stimuli were drawn from triplets composed of two real words and a nonword (e.g., *net, neck, *nep*). All the items were monosyllabic and ended with a stop consonant (a labial [/b/ or /p/], coronal [/d/ or /t/], or velar [/g/ or /k/]).

On critical trials in Experiment 1, the referent was presented along with two distractors with unrelated names and one distractor with the same initial phoneme as the referent. Figure 5 (Panel A), presents a sample display for the target word "net". Figure 5 (Panel B), presents the proportions of fixations to the target picture over time for each splicing condition. Fixations between conditions were comparable until about 600 ms after target onset, when the fixations in to the referent in the W2W1 condition began to diverge from those in the W1W1 and the NW1 conditions. The duration of the presplice fragment in the stimuli was about 400 ms, with coarticulatory cues being presumably strongest in the late portion of the vowel. Given a 200-ms delay to program and launch an eye movement, fixations occurring around 600 ms were likely to have resulted from the processing of the coarticulatory information. When this information matched an existing word, as in the W2W1 condition, fixations to the target (W1) were considerably delayed compared to when this information did not match a word, as in the N3W1 condition.

These results provide strong evidence for lexical competition. Moreover, they provide striking evidence for our claim that the general lexicon influences fixations. Both the W2W1 and the N3W1 conditions contain compa-
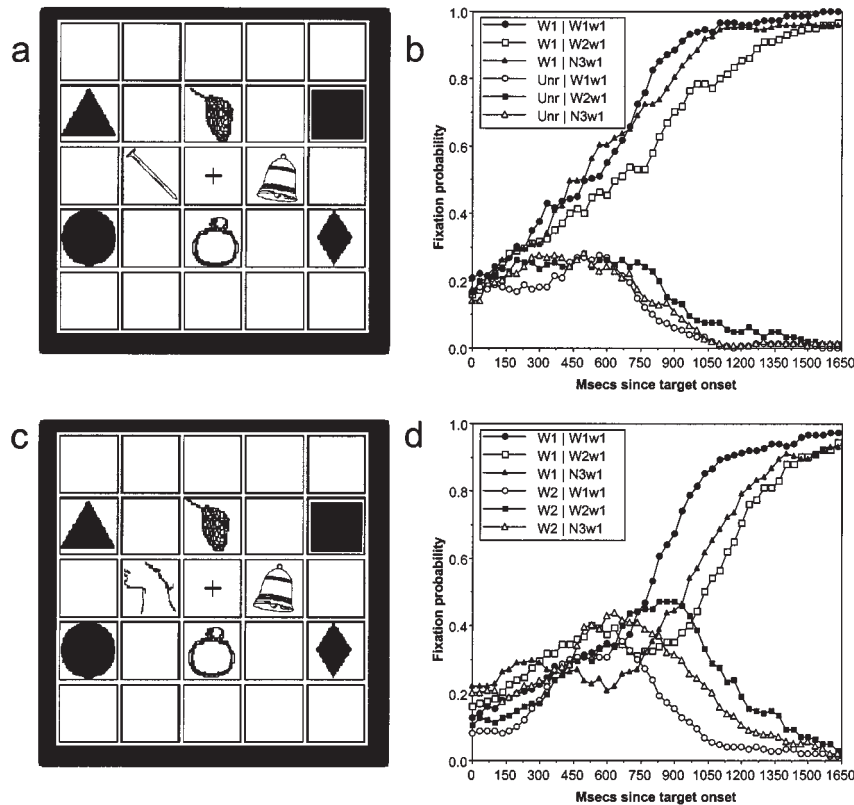
Fig. 5. Materials and results from Dahan *et al.* (in press, b). Panel a, sample stimulus display from a "cohort-absent" trial. Panel b, observed fixation probabilities for the cohort-absent condition; the word and nonword mismatch conditions yield reliable differences in target fixation probabilities. Panel c, sample stimulus display from a cohort-present trial (*net* is the target, and *neck* is the cohort). Panel d, observed fixation probabilities for the cohort-present condition.

rable subcategorical mismatches. The only difference is that in the W2W1 condition, the mismatching coarticulatory information matches a lexical competitor, whereas in the N3W1 condition, it does not match an existing word. The lexical competitor was never spoken throughout the experiment nor was it displayed in a picture. Nonetheless it had clear effects on the time course of lexical access.

In a second experiment, Dahan *et al.* (in press, b) used the same stimuli, but with displays in which both the referent and the competitor were pictured, as in Figure 5 (Panel C). This allowed us to test predictions about the time course of activation for the competitor and the target made by

models incorporating active lexical competition, such as TRACE and Shortlist (Norris, 1994). These models predict that early in the W2W1 sequence, the competitor (W2) should become highly active and compete with W1. The recognition of the target word (W1) is thus delayed. By contrast, early in the N3W1 sequence, W2 is only weakly active, so its activation has a much smaller effect on the recognition of the target word (W1). These predictions were clearly confirmed by the eye-movement results, which are presented in Figure 5, Panel D.

Fixations to the target over time indicated a fast rise of activation in the W1W1 condition, separating from the other conditions around 700 ms; the target fixations rose more slowly in the N3W1 condition, and most slowly in the W2W1. Fixations to the competitor (W2) revealed a complementary picture. The competitor picture was fixated most in the W2W1 condition, where coarticulatory information in the vowel matched the competitor's name, intermediate in the N3W1 condition, where coarticulatory information weakly matches both W1 and W2, and least in the W1W1 condition, where the coarticulatory information favors W1. In addition, simulations using TRACE provided good fits to these trends. In addition, a model of lexical decisions in which both activations from W1 and W2 can contribute to "yes" responses, predicted the pattern of lexical decisions W1W1 < N3W1 = W2W1 reported by Marslen-Wilson and Warren (1994) and McQueen *et al.* (1999).

Taken together, these three sets of studies demonstrate that: (1) the visual world paradigm provides a sensitive measure of the time course of lexical activation in continuous speech; (2) a simple and well-motivated linking hypothesis from underlying lexical activations provides an remarkably good account of the pattern and timing of fixations; and (3) the eye-movement paradigm is sensitive to general lexical variables, including effects of nondisplayed lexical neighborhoods, thus demonstrating that the instructions are being processed by the general linguistic-processing system.

GENERAL DISCUSSION

As the body of literature examining real-time language processing in natural tasks accumulates, it is becoming increasingly clear that even the earliest moments of linguistic processing are modulated by context. The relevant notion of context is likely to be a dynamic representation comprising salient entities and their properties and presuppositions shared by discourse participants, including those defined by goals, intentions, and plausible actions. Perhaps the most compelling demonstrations come from studies using head-mounted eye-tracking. We know, however, that many in the sentence-processing community remain skeptical about the conclusions

coming from this work either because they suspect that task-specific strategies might allow participants to bypass normal-language processing or because they question whether the link between eye movements and linguistic processing has been established clearly enough to interpret the results. We have addressed both of these concerns by reviewing results from a program of research in which we have used an explicit linking hypothesis to measure the time course of lexical activation in continuous speech. The results strongly support the linking hypothesis and clearly demonstrate effects that cannot be attributed to task-specific strategies. We suspect, though, that a set of deeper beliefs might also underlie the skepticism of some in the community. Some of these beliefs have been articulated most forcefully and eloquently by Fodor (1983).

One of these biases is that one cannot study context coherently because it involves general-purpose cognitive abilities, which on this view are seen as too unconstrained to be scientifically tractable. However, researchers in the psycholinguistics, computational, and linguistics communities continue to make progress defining and formalizing mechanistic theories of relevant context and how it is related to linguistic form. The second bias is that the goal of real-time sentence-processing research should be to determine how listeners compute context-independent representations. The output of this "input" system will be a linguistic representation that can be computed quickly, because it is encapsulated. Moreover, it will be general enough to serve as the basis of more context-specific representations that will be computed by the more general cognitive system. However, we have seen that linguistic input can be immediately integrated with nonlinguistic context. Thus, the speed of real-time comprehension cannot form the basis of an argument for encapsulation. Moreover, an increasing body of research on the visual system—perhaps the strongest *a priori* candidate for an input system— suggests that processing in even the earliest levels of the visual cortex is modulated by task-specific constraints (e.g., Gilbert, 1998; Gottlieb, Kusunoki, & Goldberg, 1998). Results such as these are calling into question the longstanding belief that perceptual systems create context-independent perceptual representations.

REFERENCES

Allopenna, P., 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. T. M. (1998). Ambiguity in sentence processing. *Trends in Cognitive Sciences, 2,* 146–152.

Altmann, G. T. M., Haywood, S., & Kamide, Y. (2000). Anticipating grammatical function:

Evidence from eye movements. Paper presented at the *13th Annual CUNY Conference on Human Sentence Processing,* La Jolla, California.

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

Arnold, J. E., Eisenband, J. G., Brown-Schmidt, S., & Trueswell, J. C. (2000). The rapid use of gender information: Evidence of the time course of pronoun resolution from eyetracking. *Cognition, 76,* B13–B26.

Chambers, C. G., Tanenhaus, M. K., & Magnuson, J. S. (2000). Does real-world knowledge modulate referential effects on PP-attachment? Evidence from eye movements in spoken language comprehension. Paper presented at the *13th Annual CUNY Conference on Human Sentence Processing,* La Jolla, California.

Clark, H. H. (1992). *Arenas of language use.* Chicago, Illinois: University of Chicago Press.

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,* 84–107.

Crain, S., & Steedman, M. J. (1985). On not being led up the garden path. In D. Dowty, L. Kartunnen & A. Zwicky (Eds.), *Natural language parsing* (pp. 320–358). Cambridge: Cambridge University Press.

Dahan, D., Magnuson, J. S., & Tanenhaus, M. K. Time course of frequency effects in spoken-word recognition: evidence from eye movements, submitted.

Dahan, D., Magnuson, J. S., Tanenhaus, M. K., & Hogan, E. Subcategorical mismatches and the time course of lexical access: Evidence for lexical competition. *Language and Cognitive Processes,* in press.

Dahan, D., Swingley, D., Tanenhaus, M. K., & Magnuson, J. S. (2000). Linguistic gender and spoken-word recognition in French. *Journal of Memory and Language, 42,* 465–480.

Eberhard, K. M. (1998). Watching speakers speak: Using eye movements to study language production. Invited paper presented at the *70th Annual Meeting of the Midwestern Psychological Association,* Chicago, Illinois.

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

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

Fodor, J. A. (1983). *Modularity of mind.* Cambridge, Massachusetts: MIT Press.

Francis, W. N., & Kucera, H (1982). *Frequency analysis of English usage: Lexicon and grammar.* Boston, Massachusetts: Houghton Mifflin Company.

Gaskell, M. G., & Marslen-Wilson, W. D. (1997). Integrating form and meaning: A distributed model of speech perception. *Language and Cognitive Processes, 12,* 613–656.

Gibson E., & Pearlmutter, N. (1998). Constraints on sentence comprehension. *Trends in Cognitive Science, 2,* 262–268.

Gilbert, C. D. (1998). Adult cortical dynamics. *Physiological Review, 78,* 467–485.

Gottlieb, J., Kusunoki, M., & Goldberg, M. E. (1998). The representation of visual salience in monkey posterior parietal cortex. *Nature, (London) 391,* 481–484.

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

Hayhoe, M. (2000). Vision using routines: A functional account of vision. *Visual Cognition, 7,* 43–64.

Hurewitz, F., Brown-Schmidt, S., Gleitman, L., & Trueswell, J. C. (2000). When production precedes comprehension: Children fail to understand constructions that they freely and accu-

rately produce. Paper presented at the *13th Annual CUNY Conference on Human Sentence Processing,* La Jolla, California.

Kako. E., & Trueswell, J. C. (2000). Mapping referential competition and the rapid use of verb semantic constraints. Poster presented at the *13th Annual CUNY Conference on Human Sentence Processing,* La Jolla, California.

Luce, R. D. (1959). *Individual choice behavior.* New York: Wiley.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The Neighborhood Activation Model. *Ear & Hearing, 19,* 1–36.

Magnuson, J. S., Tanenhaus, M. K., Aslin, R. N., & Dahan, D. (1999). Spoken word recognition in the visual world paradigm reflects the structure of the entire lexicon. In M. Hahn & S. Stoness (Eds.), *Proceedings of the Twenty First Annual Conference of the Cognitive Science Society,* (pp. 331–336). Mahwah, NJ: Erlbaum.

Marslen-Wilson, W. (1987). Functional parallelism in spoken word-recognition. *Cognition, 25,* 71–102.

Marslen-Wilson, W. (1990). Activation, competition, and frequency in lexical access. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing. Psycholinguistic and computational perspectives* (pp. 148–172). Hove, UK: Erlbaum.

Marslen-Wilson, W. (1993). Issues of process and representation in lexical access. In G. T. M. Altmann, and R. Shillcock (Eds.), *Cognitive models of speech processing: The Second Sperlonga Meeting.* (pp. 187–210). Hove, England UK: Lawrence Erlbaum Associates.

Marslen-Wilson, W., & Warren, P. (1994). Levels of perceptual representation and process in lexical access: Words, phonemes, and features. *Psychological Review, 101,* 653–675.

Matin, E., Shao, K., & Boff, K. (1993). Saccadic overhead: information processing time with and without saccades. *Perception & Psychophysics, 53,* 372–380.

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

McQueen, J. M., Norris, D., & Cutler, A. (1999). Lexical influence in phonetic decision making: evidence from subcategorical mismatches. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 1363–1389.

Meyer, A. S., Sleiderink, A. M., & Levelt, W. J. M. (1999). Viewing and naming objects: eye movements during noun phrase production. *Cognition, 66,* B25–B33.

Nadig, A., & Sedivy, J. C. (2000). Children's use of referential pragmatic constraints in production and processing. Paper presented at the *13th CUNY Conference on Human Sentence Processing,* La Jolla, California.

Nicol, J., & Swinney, D. (1989). The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research, 18,* 5–19.

Norris, D. (1994). Shortlist: a connectionist model of continuous speech recognition. *Cognition, 52,* 189–234.

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

Pollack, I., Rubenstein, H., & Decker, L. (1959). Intelligibility of known and unknown message sets. *Journal of Acoustical Society of America, 31*, 273–279.

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

Runner, J. T., Sussman, R., & Tanenhaus, M. K. (2000). Binding reflexives and pronouns in real-time processing. Poster presented at the *13th Annual CUNY Conference on Human Sentence Processing,* La Jolla, California.

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

Simons, D. J. (2000). Change blindness and visual memory. *Visual Cognition,* [S.I] *7,* 1–416.

Sommers, M. S., Kirk, K. I., & Pisoni, D. B. (1997). Some consideration in evaluating spoken word recognition by normal-hearing, noise-masked normal hearing, and cochlear implant listeners. I: The effects of response format. *Ear and Hearing, 18,* 89–99.

Spivey-Knowlton, M. J. (1996). Integration of visual and linguistic information: Human data and model simulations. Ph.D. dissertation, University of Rochester.

Spivey, M. J., & Marian, V. (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science, 10,* 281–284.

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. K., Eberhard, K. E., & Sedivy, J. C. (2000). Eye movements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution. *Cognitive Psychology,* in press.

Streeter, L. A., & Nigro, G. N. (1979). The role of medial consonant transitions in word perception. *The Journal of Acoustical Society of America, 65,* 1533–1541.

Sussman, R., & Sedivy, J. C. (2000). Using eyetracking to detect and describe failed gap effects. Poster presented at the *13th Annual CUNY Conference on Human Sentence Processing,* La Jolla, California.

Tanenhaus, M. K., Dahan, D., Magnuson, J. S., & Hogan, E. (2000). Tracking the time course of subcategorical mismatches on lexical access. Paper presented at the *13th Annual CUNY Conference on Human Sentence Processing,* La Jolla, California.

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., Eberhard, K. M., & Sedivy, J. C. (1996). Using eye movements to study spoken language comprehension: Evidence for visually mediated incremental interpretation. In T. Inui & J. McClelland (Eds.), *Attention & performance XVI: Integration in perception and communication.* (pp. 457–478). Cambridge, Massachusetts: MIT Press.

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

Tanenhaus, M. K., & Trueswell, J. (1995). Sentence comprehension. In J. Miller & P. Eimas (Eds.), *Speech, language, and communication* (pp. 217–595). San Diego, California: Academic Press.

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

Whalen, D. H. (1984). Subcategorical phonetic mismatches slow phonetic judgments. *Perception & Psychophysics, 35,* 49–64.

Whalen, D. H. (1991). Subcategorical phonetic mismatches and lexical access. *Perception & Psychophysics, 50,* 351–360.

Yee, E., Blumstein, S., & Sedivy, J. C. (2000). The time course of lexical activation in Broca's aphasia: Evidence from eye movements. Poster presented at the *13th Annual CUNY Conference on Human Sentence Processing,* La Jolla, California.