ELSEVIER

Cognitive Science 28 (2004) 721–749

COGNITIVE SCIENCE

http://www.elsevier.com/locate/cogsci

Disfluencies, language comprehension, and Tree Adjoining Grammars

Fernanda Ferreira*, Ellen F. Lau, Karl G.D. Bailey

Department of Psychology, Michigan State University, East Lansing, MI 488424-1117, USA

Received 30 May 2003; received in revised form 16 October 2003; accepted 28 October 2003

Available online 19 March 2004

Abstract

Disfluencies include editing terms such as *uh* and *um* as well as repeats and revisions. Little is known about how disfluencies are processed, and there has been next to no research focused on the way that disfluencies affect structure-building operations during comprehension. We review major findings from both computational linguistics and psycholinguistics, and then we summarize the results of our own work which centers on how the parser behaves when it encounters a disfluency. We describe some new research showing that information associated with misarticulated verbs lingers, and which adds to the large body of data on the critical influence of verb argument structures on sentence comprehension. The paper also presents a model of disfluency processing. The parser uses a Tree Adjoining Grammar to build phrase structure. In this approach, filled and unfilled pauses affect the timing of Substitution operations. Repairs and corrections are handled by a mechanism we term "Overlay," which allows the parser to overwrite an undesired tree with the appropriate, correct tree. This model of disfluency processing highlights the need for the parser to sometimes coordinate the mechanisms that perform garden-path reanalysis with those that do disfluency repair. The research program as a whole demonstrates that it is possible to study disfluencies systematically and to learn how the parser handles filler material and mistakes. It also showcases the power of Tree Adjoining Grammars, a formalism developed by Aravind Joshi which has yielded results in many different areas of linguistics and cognitive science.

Keywords: Parsing; Disfluencies; TAG; Syntax

* Corresponding author.
E-mail address: Fernanda@eyelab.msu.edu (F. Ferreira).

0364-0213/$ – see front matter © 2004 Cognitive Science Society, Inc. All rights reserved.
doi:10.1016/j.cogsci.2003.10.006

1. Introduction

As psycholinguists who work on the question of how people interpret sentences, our research has been informed by decades of work treating spoken and written language in essentially the same way. Clearly the input modalities for listening and reading are different, but investigations of syntactic analysis in particular have assumed that the material the parser receives is the same once the system has performed word recognition and extracted whatever prosodic or graphic information is in the stimulus. Reading has been studied much more extensively than spoken language processing, in part because it is much easier to present materials visually to experimental participants, and perhaps more importantly, because the best online measures for examining moment-by-moment processing of language rely on visual presentation—the monitoring of eye fixations, for example. Only recently have techniques been developed for studying spoken language comprehension that have anything approximating the same degree of resolution. For example, the auditory moving window technique (Caplan & Waters, 2001; Ferreira, Anes, & Horine, 1996; Ferreira, Henderson, Anes, Weeks, & McFarlane, 1996) and the visual world paradigm (Keysar, Barr, Balin, & Brauner, 2000; Spivey, Tanenhaus, Eberhard, & Sedivy, 2002; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy 1995; see also Henderson & Ferreira, 2004) allow psycholinguists to measure aspects of the language system's performance as an utterance unfolds. Still, these paradigms have many drawbacks compared to the unobtrusive monitoring of eye movements during silent reading.

Many of the conclusions that emerged from studies of reading have turned out to generalize to spoken language. For instance, it appears that language comprehension in both modalities is incremental (Altmann & Kamide, 1999; Just & Carpenter, 1992; Tabor, Juliano, & Tanenhaus, 1997; Tanenhaus, Carlson, & Trueswell, 1989). In addition, points of syntactic complexity and choice slow down processing and cause the parser to search for information that can help it select from among competing structures (Ferreira & Clifton, 1986; Frazier & Rayner, 1982; MacDonald, Pearlmutter, & Seidenberg, 1994). And most relevant for the issues that will be discussed in this paper, information about verb argument structure is critical for all aspects of interpretation (Boland, Tanenhaus, & Garnsey, 1990; MacDonald, 1994; Mitchell & Holmes, 1985; Rayner, Carlson, & Frazier, 1983).

Yet there are major differences between spoken and written language. For example, when listeners process speech they create prosodic representations based on acoustic cues in the input, and there is evidence that prosodic constituency affects the way that syntactic (Kjelgaard & Speer, 1999) and semantic (Schafer, Carlson, Clifton, & Frazier, 2000) representations are constructed. (It should also be noted that many researchers believe prosodic representations are created during reading as well—for example, see Fodor (1998) and Bader (1998)—but obviously they are constructed quite differently and their characteristics can be difficult to observe.) An even starker contrast between listening and reading is that speech produced in natural contexts contains material that we will refer to collectively as **disfluencies**. For example, consider *Bill said you should drop the frog*. A long silent pause before *drop* might occur, or a "filled" pause such as *uh* or *um* (*Bill said you should uh drop the frog*). Or, a word might be repeated: *Bill said you should drop the frog*. These two kinds of disfluencies—silent/filled pauses[1] and repeats—are usually attempts by the speaker to give him- or herself more time to plan (e.g., Clark & Wasow, 1998). Another type of disfluency, and one that will be the focus of this paper, is **revisions**, as in *Bill said you should put—drop the ball*. Revisions are not due to the speaker trying to buy more time to prepare; indeed, as Shriberg (2001) observed, speakers who produce a large number of revisions and other sorts of repairs are fast talkers who probably underestimate the amount of time they need to plan their utterance compared to those whose speech is peppered with filled pauses and repeats. Revisions, then, are errors that speakers make when they fail to plan to a sufficient degree.

Disfluencies are important phenomena for psycholinguists to investigate, for at least three reasons. First, if our goal is to develop theories of comprehension that are general and complete, then we cannot ignore a phenomenon as universal and widespread as disfluencies. It is estimated that for every 100 words of speech there will be at least six disfluencies (Bortfeld, Leon, Bloom, Schober, and Brennan, 2001; Fox Tree, 1995), making them almost as much of a challenge for the comprehension system as syntactic ambiguity (which, needless to say, has been the subject of hundreds of psycholinguistic investigations). It is clear that the parser, which presumably evolved to deal with naturalistic, interactive conversations, must have developed mechanisms for handling this aspect of the input. Yet disfluencies have been almost entirely ignored in models of human sentence processing (but not in computational linguistics; more on this later), perhaps because it is at least tacitly assumed that the parser must somehow filter the disfluencies so that it can then operate on just words and prosodically licensed acoustic cues. But this assumption is surely false. One of the few issues on which there is almost universal agreement in the field of sentence comprehension is incrementality: The system interprets input as it is received (Marslen-Wilson & Tyler, 1981). Indeed, not only does the processor extract as much information as it can from each successive bit of input as it is encountered, it even tries to predict upcoming words (Altmann & Kamide, 1999). The problem is that if the system is incremental, then the filtering view cannot be right, because filtering would require input to be buffered, disfluencies to be removed, and the parser and other components of the processor to do their work of building syntactic structures and interpretations only afterwards. If we abandon this idea that utterances are interpreted after disfluencies have been filtered (as we surely must), then we find ourselves with the task of having to explain how they are handled. And that is a question that has received almost no attention in the field of psycholinguistics. Our recent research attempts to remedy this situation.

Second, disfluencies are important to investigate because they challenge the parser in important ways. We will provide a few examples here. Consider a highly disruptive type of disfluency: revisions, as in *Bill said you will put—you should drop the ball*. If the parser operated without considering the possibility of disfluencies, it would construct a complement clause with two sequences of a subject, a modal, and a main verb. But of course the parser cannot do this if it uses the grammar to create structural representations—that is, if there is transparency (Berwick & Weinberg, 1984) between the grammar and the parser—because this string is illicit. At a descriptive level, we know that what the parser somehow does is to integrate *you will put* as part of the complement clause, but at some point it might use the ungrammaticality of the resulting sequence *you will put you should drop the* as evidence for the presence of a disfluency (as Core & Schubert, 1999 propose for computer speech parsing). The sequence *you will put* would then be expunged from the structure, leaving just *you should drop*. In the second-last section of this paper, we describe a model of “disfluency reanalysis” that describes how some of these operations might work. This approach uses a Tree Adjoining Grammar (TAG)-based parser to build syntactic structure for utterances containing disfluencies, including filled pauses, repeats, and revisions.

A filled pause might be the easiest disfluency for the parser to deal with, because it is clearly marked as such. Moreover, if a sentence contains only an *uh* or *um*, then the parser does not have the problem of coping with lexical material that does not belong in the sentence's final representation. But this does not imply that filled pauses are without effect, and indeed in our work (Bailey & Ferreira, 2001, 2003a, 2003b, in press) we postulated that they might influence the parser in two ways: First, during the reception of a filled pause, the parser might be in a state of waiting for the *uh* or the *um* to end and for words to resume coming in; that state might have consequences for the structure or structures under construction in memory. A second possible effect is that such disfluency terms might even serve as helpful cues to the structure of the utterance at points of ambiguity, thus revealing that some disfluencies might actually be useful to the parser. We will elaborate on both points in this article.

Revisions, as discussed earlier, challenge the parser because the reparandum (the portion that is wrong) must be edited out of the structure and replaced with the corrected material. Some revisions include repetition of unchanged material in addition to the corrected material, as in *Bill said I—Bill said you should drop the ball*. Other revisions only involve substitution of the changed material, as in *Bill said I—you should drop the ball*. In this paper we will refer to this particular form of revision as **replacement**. In our work we have chosen to examine a simple type of one word replacement: verb replacements, as in *you should put–drop the frog*. This kind of example is of interest to us because it allows us to investigate a critical question in sentence processing: how verb information is used during parsing. Lexically based approaches to parsing (such as Lexicalized-TAG or L-TAG) assume that structures are directly associated with lexical items, and the type of word that carries the richest structural information is the verb (McElree & Griffith, 1995; Shapiro, Zurif, & Grimshaw, 1987; Trueswell & Kim, 1998). Thus, an utterance such as *You should put–drop the frog* presents the parser with an interesting problem: The verb *put* is initially retrieved along with its associated argument structure, because of course the parser does not yet know that this word is the reparandum. Some work in computational linguistics does suggest that reparanda might be signaled by acoustic cues (see Nakatani & Hirschberg, 1994 for review), but the information tends to occur late in the reparandum, and little is known about the reliability of such cues or the ability of the parser to use this information; therefore, we will ignore this interesting issue for the purposes of this paper. When the correct verb *drop* arrives, it is retrieved from the lexicon together with its argument structure. How are the two verb argument structures reconciled? Is the one associated with *put* entirely erased, or do its effects persist? The study we report in this paper addresses this question directly.

We have argued that disfluencies belong in a theory of spoken language comprehension because they occur in a large number of the utterances the processor actually encounters, and because they present interesting computational challenges for the parsing mechanism. Our final reason for extending sentence comprehension to the study of disfluencies is that if the parser turns out to have robust mechanisms for handling them, then it will be clear that models of structure-building and interpretation must take spoken language much more seriously than they have up till now. This argument is based on the obvious fact that *only* spoken language contains disfluencies—disfluencies in written texts are inserted for effect, as in *let's talk about* *your uh problem*. And written texts virtually never contain repetitions or replacements (except for text drafts that have not yet been properly edited). Moreover, disfluencies are integral to speech: If a person is talking, then he or she is producing disfluencies (every 15 or so words, according to estimates obtained from large corpora; Bortfeld et al., 2001; Fox Tree, 1995).

2. Previous treatments of disfluencies

2.1. Disfluencies in computational linguistics

For automatic speech recognizers to function at an adequate level, they must be capable of handling disfluencies. Yet the best way of achieving this goal is a matter still under debate. While some researchers have proposed filtering to try to remove all disfluencies from the input before parsing begins (Charniak & Johnson, 2001; Stolcke & Shriberg, 1996), others have argued that information in a reparandum (for instance) might have to be preserved for later stages in the parser, such as for reference resolution (Core & Schubert, 1999, who use the real example *Take the oranges to Elmira um I mean take them to Corning* to illustrate this point). But regardless of how disfluency information is treated relative to the fluent content, all models have to deal with the same initial problem: how to recognize disfluencies in the first place.

A wide variety of approaches have been taken to solving the recognition problem. Some use grammars or meta-rules (rules that operate on grammatical structures output by a parser rather than on words) that can account for disfluencies (McKelvie, 1998; Heeman & Allen, 1997); others rely purely on statistical co-occurrences in N-gram type models (Stolcke & Shriberg, 1996). One reason why so many widely differing models exist and why none of them locate disfluencies with good accuracy is that disfluencies comprise a very heterogeneous class. Filled pauses, for instance, are a relatively closed set and may be more amenable to co-occurrence modeling (Stolcke & Shriberg, 1996). But other disfluencies such as repetitions, revisions, and deletions are much more difficult to handle statistically because they have no consistent lexical form, and they vary widely in span (from as few as two words to entire sentences, with the mode being just two words; Shriberg, 1996). A related issue is that, while disfluent speech is common enough that the parser needs some way of dealing with it, the number of fluent speech occurrences is much greater than disfluent speech occurrences if summed across all word boundaries. This can make it difficult for a purely probabilistic system to pull out features that are specifically associated with disfluencies (Stolcke et al., 1998).

Although a great deal of progress still needs to be made before speech processors can detect and handle disfluencies automatically, much of what we know about disfluencies and their distributional characteristics is due to the efforts of computational linguists who have carefully examined large speech corpora such as SWITCHBOARD (Godfrey, Holliman, & McDaniel, 1992), TRAINS (Heeman & Allen, 1995), AMEX (Kowtko & Price, 1989), and ATIS (MADCOW, 1992). The following are some of the facts that have emerged: First, as mentioned earlier, disfluencies occur at the rate of about 6 per 100 words (Bortfeld et al., 2001; Fox Tree, 1995). In human–computer interactions they are much rarer (occurring only about 1% of the time), possibly because people prepare much more carefully before speaking to a machine (Oviatt, 1995; Shriberg, 1996). Indeed, often the devices are designed so that speakers push a button when they are ready to talk to the computer. Second, the most common type of disfluency is the filled pause, followed by word repetitions. All other types are comparatively rare (see Shriberg, 1996, for frequency values). A third generalization about disfluencies based on examination of corpora is that disfluencies are most likely to occur at the beginning of a sentence (Shriberg, 1996), and repeats in particular are frequent right at the start or one word into a syntactically complex phrase or clause (Clark & Wasow, 1998). Also, the longer a sentence, the more likely a disfluency is to occur.

It is also of interest that individuals vary a great deal in how disfluent they are: Some people can get by with fewer than one disfluency per 100 words, while others produce 10 or more. Moreover, people differ in the type of disfluency they tend to rely on. Slow speakers tend to use empty and filled pauses as well as repeats; fast speakers produce a large number of deletions, revisions, and outright errors (Shriberg, 1996). Clearly, then, some people plan their speech carefully and others formulate more “on the fly.” This generalization might seem fairly mundane, but it does have important implications for models of sentence production, which have assumed that the system is either incremental (planning and articulation occur in parallel) or not. These findings concerning individual differences in disfluencies suggest that degree of incrementality must be viewed as something like a personality variable. Moreover, incrementality in production is often invoked to explain the apparent fluency of speech. But the data from large corpora suggest that (a) speakers are not all that fluent, and (b) incrementality does not guarantee fluency, and in fact lack of planning can lead a speaker to produce the most disruptive types of disfluencies, corrections and reformulations.

To sum up, disfluencies are obviously an important problem for the field of computational linguistics, because no automatic recognition system will be successful unless disfluencies can be detected and processed properly. In dealing with revisions and repetitions, the most challenging type to recognize and process, recent systems have only managed to reduce the misclassification rate on such disfluencies from a null rate of 5–6% to an error rate of 2–3% (Charniak & Johnson, 2001; Stolcke et al., 1998). Nevertheless, basic research on the characteristics of disfluencies in natural conversations (between humans and between humans and machines) has yielded a great deal of useful information, and more work along these lines clearly needs to be conducted.

2.2. Disfluencies in psycholinguistics

While much work on disfluencies has been concerned with the problems that speakers face in monitoring for and correcting errors in their own speech (Blackmer & Mitton, 1991; Brédart, 1991; Levelt, 1983; van Wijk & Kempen, 1987), there have been several studies focusing on disfluencies in language comprehension. Some of these have examined the effects of the mere presence of disfluencies in utterances on decisions and judgments made by listeners. Arnold, Fagnano, and Tanenhaus (2003) found that filled pauses in an instruction to move a particular object increased the likelihood that subjects would look to an object that had not been previously mentioned in the discourse. When the instruction was fluent, on the other hand, subjects were more likely to look at an object that had already been mentioned before. It appears, then, that comprehenders use a disfluency as a source of information for segmenting the visual world under discussion into given and new elements (Clark & Haviland, 1977). Listeners know that a speaker is more likely to have trouble articulating the name of an object not yet established in the discourse compared to one that has already been mentioned.

Disfluencies also reflect on the speaker's state of knowledge in systematic ways, and listeners use this information during comprehension. Brennan and Williams (1995) had people listen to answers to questions that had been posed earlier. Some of the answers included a filled pause and others did not; in addition, the answer was sometimes right and sometimes wrong. These researchers found that the disfluencies in the answer made subjects judge that speakers were less sure about their correct answers, but also more likely to actually know the answer to the questions that they claimed they could not answer.

Disfluencies that involve repetition, revision, or abandonment of utterances have more serious effects on processing. Levelt (1983, 1984) referred to the issue of dealing with repeated or repaired material as the listener's continuation problem. When faced with material that must be excised in order to form an ideal utterance (Clark & Clark, 1977), the listener must determine where in the utterance the material to be excised (i.e., the reparandum) begins and where it ends, and he or she must also locate the start of the repair. At any disfluency, the listener must decide between four possible options (Clark, 1996). First, the speaker may be restarting by replacing the immediately preceding word. Second, the speaker may be replacing material, but may be retracing from an earlier point in the utterance. Third, the preceding utterance may have been abandoned, and the speaker may have decided to start anew. Finally, the disfluency may be what Levelt and others have referred to as a covert repair, that is a disfluency where no replacement or repetitions of material occurs (e.g., a filled or unfilled (but not prosodic) pause). Levelt (1983, 1984) suggested two conventions that would allow listeners to solve this continuation problem.

The first convention, Word-Identity, states that when some word in the original utterance is identical to the first word in the repair, the repair is considered to be a continuation of the utterance starting with that word. If more than one instance of the word occurs in the original utterance, the last instance is taken to be the appropriate one. A second convention, Category-Identity, is meant to apply in those cases where the first word in the repair is not identical to any in the original utterance. In such cases, the last word in the original utterance that matches the syntactic category of the first word in the repair is taken to be the starting point of the continuation.

Of course, in order to take advantage of these conventions, the language comprehension system would need to be able to identify disfluencies easily. Ideally, speakers would mark a disfluency using pausing and prosodic changes. When such markers are present, listeners judge utterances to be easier to comprehend, and are able to initiate a repetition more quickly (Howell & Young, 1991). However, such cues are usually absent (Cutler, 1983; Levelt & Cutler, 1983). Yet even without reliable prosodic cues, evidence from word gating experiments suggest that subjects are able to identify that a disfluency has occurred even before the end of the first word in the repair (Lickley & Bard, 1998). This suggests that the language comprehension system is able to identify a disfluency, likely through the use of a combination of cues (in some manner that is as yet not understood). Notice, however, that the information comes too late to allow a comprehender to know *at the reparandum* that material will soon be overwritten. It is the repair portion that contains useful cues, and of course the repair follows the portion that was spoken in error.

Given that the language comprehension system can detect disfluencies, what effects might they have on the processing of an utterance? Unfortunately, there is almost no evidence available yet to answer this question (a situation which partly inspired the line of research we described in this article). Levelt's description of the listener's continuation problem suggests that disfluencies add significantly to the processing system's computational burden. Clearly, though, some disfluencies are easier to handle than others; for example, it seems likely that repetitions are easier to deal with than revisions, given that a repetition would invoke the Word-Identity convention and would only require a search back to the preceding word, while a revision might require a search further back (if the Word-Identity convention is invoked) or a switch to the Category-Identity convention. A study conducted by Fox Tree (1995) using a word monitoring paradigm supports this prediction. When monitoring for a word that followed a disfluency, subjects were slower to respond when the disfluency involved revision compared with repetitions.

Brennan and Schober (2001) also examined the processing of repairs, but obtained somewhat different results. Whereas Fox Tree (1995) found increased latencies for revision disfluencies, Brennan and Schober found that revisions (specifically, replacements) did not lead to an increased latency to respond relative to fluent controls in a paradigm where subjects selected an appropriate colored shape from a display in response to an instruction. An examination of subjects' errors suggested that the amount of misleading information that listeners heard affected the likelihood that they would make a commitment to the incorrect response. Thus, subjects made fewer errors when the incorrect word was only partially articulated compared to a condition in which the wrong word was spoken in its entirety; and still fewer errors occurred when the erroneous word was interrupted and followed by a filled pause (e.g., *uh*).

In summary, psycholinguistic research has yielded several important findings concerning the ways that disfluencies are handled during high-level comprehension. First, it is clear that disfluencies have both early and late effects on processing. Disfluencies can influence decisions immediately, as with Arnold et al.'s (2003) demonstration that the presence of disfluency affected the immediate fixation of objects; or disfluencies can operate at the level of meta-judgments, as in the studies demonstrating the way disfluencies affect judgments of speakers' knowledge and understanding of a domain in which they are being questioned (Brennan & Williams, 1995; Smith & Clark, 1993). When considering disfluencies that introduce material that must be deleted in order to form a ideal grammatical utterance, psycholinguists have suggested that their structure may play a part in solving the listener's continuation problem. Disfluencies may be produced according to certain conventions (Levelt, 1983, 1984) that listeners can take advantage of in order to solve the continuation problem and may have features associated with them that allow for immediate detection (Howell & Young, 1991; Lickley & Bard, 1998). There is some evidence that replacements are harder to deal with than repetitions (Fox Tree, 1995). Other evidence, however, suggests that the language processor is sensitive to the amount of information contained in a disfluency; if a replacement is produced so that the wrong word is only partially articulated and is then followed by an *uh*, it is as easy for the system to handle as a fluent utterance (Brennan & Schober, 2001). Of course, we are referring only to aspects of the comprehension system that can be measured when participants listen to simple commands to move objects; we do not know yet whether the parser which builds the tree on which the interpretation is built had more trouble handling the replacement compared to a repair.

3. Our work on disfluencies in parsing

3.1. Filled pauses

As the brief review we presented in the preceding section should make clear, until we began our line of research, almost no psycholinguistic work on disfluencies focused on syntactic operations. But it seems obvious that disfluencies play some role in parsing, and so for the last few years we have been examining how disfluencies might affect the process of building a syntactic structure for a sentence (Bailey & Ferreira, 2001, 2003a, 2003b, in press). We began with the most conservative hypothesis: At the very least, disfluencies must influence parsing because as they unfold the parser has to sit in a holding pattern until it can once again go back to building its phrase marker. Therefore, disfluencies might affect temporal aspects of structure building. Imagine that a listener encounters *While the man hunted the deer* . . . and then hears a disfluency containing a filler or two (i.e., *While the man hunted the deer uh uh* . . .). Assume that the listener has analyzed *the deer* as the object of *hunted* (perhaps because that analysis is more frequent or is syntactically simpler). If time influences the "strength" of a syntactic analysis, then the disfluencies will serve to reinforce that interpretation, making reanalysis more difficult if it should be required. The Head Position Effect (HPE) found by Ferreira and Henderson (1991a) is consistent with the idea that time would work in just this way. Ferreira and Henderson (1991a) demonstrated that reanalysis of a sentence such as (1a) is much harder when the ambiguous NP is lengthened with postnominal modifiers (as in 1c) than when it is lengthened with prenominal adjectives (as in 1b). (The underlined word is the one that disambiguates the sentence.) Ferreira and Henderson explained the effect by suggesting that the sentence comprehension system assigns thematic roles to syntactic heads, and the longer it has been committed to the wrong thematic analysis, the harder the structure is to reanalyze.

(1)	a. ***Baseline*** (*Plain Ambiguous NP*):	While the man hunted the deer <u>ran</u> into the woods.
	b. *Head late, words*:	While the man hunted the brown and furry deer <u>ran</u> into the woods.
	c. *Head early, words*:	While the man hunted the deer that was furry <u>ran</u> into the woods.
	d. *Head late, disfluencies*:	While the man hunted the uh uh deer <u>ran</u> into the woods.
	e. *Head early, disfluencies*:	While the man hunted the deer uh uh <u>ran</u> into the woods.

The HPE gave us our first idea about the possible influence of disfluencies on comprehension (Bailey & Ferreira, 2003a, 2003b, in press). We speculated that disfluencies might influence syntactic analysis and reanalysis processes because their presence would affect how long the parser is committed to a particular structure. This finding would be significant, because it would be the first to suggest that disfluencies systematically influence the parser's syntactic procedures. To test this possibility, we designed a study with the logic illustrated in (1). The Baseline condition measures the difficulty of t he garden-path sentence without any "extra" material. The two conditions containing words allowed us to make sure we could replicate the HPE (1b and 1c). The most important manipulation, of course, is the presence of disfluencies in the early position (1e) and in the late position (1d). All of the sentences were spoken as naturally as possible, and the disfluencies were "staged"—that is, the speaker (KGDB) produced the sentences with the disfluencies in the intended locations, making the renditions sound as natural as possible. During post-experiment debriefing sessions the participants indicated that the sentences sounded like they were naturally produced.

The participants' main task was to listen to sentences and then push a button to indicate whether they thought they were acceptable in English. This task is used extensively in studies of parsing and reanalysis (e.g., Devescovi et al., 1997; Grossman, Mickanin, Onishi, & Hughes, 1996; McElree & Griffith, 1995). The logic is that if the parser fails to find a good analysis for an utterance, then to the parser the sentence is ungrammatical. Therefore, the subject will label it as such in the experiment (see Ferreira & Henderson, 1991b, for a more detailed discussion of the logic). Participants were explicitly told that their grammaticality decisions should not be based solely on the presence or absence of a disfluency, and they were told simply to expect some sentences similar to "real" utterances in that they would contain "uh"s. The participants ($n = 30$) judged the grammaticality of a variety of sentences representing the five conditions illustrated in (1), and they evaluated a large number of grammatical and ungrammatical fillers as well. The results are shown in Table 1.

As the table clearly shows, the disfluencies behaved exactly the same way that words did. Our explanation of this finding is that disfluencies affect parsing for the reasons given by Ferreira and Henderson (1991b): Thematic roles are assigned at heads, and the longer the sentence comprehension system has been committed to the wrong analysis, the more difficult it is to revise that analysis. On this view, the critical element is **time** since the thematic role was assigned. Disfluencies influence comprehension, then, because they shorten or lengthen the amount of time that the parser is committed to its syntactic analyses. When the parser encounters *While the man hunted the deer uh uh ran into the woods*, it first makes *the deer* the object of *hunt* (both because of general parsing principles such as late closure (Frazier & Rayner, 1982), and also because *hunt* is strongly transitive). The parser's commitment to that transitive analysis of the subordinate clause persists while *uh uh* is being processed, just as

Table 1
Percentage of sentences judged grammatical for each of the conditions represented in (1)

Condition	%
Baseline	83
Head late, words	80
Head late, disfluencies	85
Head early, words	59
Head early, disfluencies	60

it does when a postnominal modifier occurs in that same sentential position. As a result, the reanalysis forced by the presence of *ran* is difficult because the parser has been committed to the incorrect analysis longer than it would have been if the disfluencies had not occurred. Perhaps this occurs because the structure that the parser had been entertaining gains strength over time; or perhaps the structure that is initially not chosen but is ultimately correct (the intransitive one) loses activation and therefore is difficult to retrieve; or certainly both mechanisms could operate at the same time. The important point, though, is that disfluencies influence the parser because they change how long it is committed to the structure that is under construction when the disfluency is encountered.

This is one of the ways we believe disfluencies affect parsing. Another is that disfluencies might help the parser resolve syntactic ambiguity. This idea stems from the observation that disfluencies have a particular distribution with respect to the syntactic constituents of an sentence. For example, disfluencies are more likely to occur at the start of a major constituent such as a clause (Clark & Wasow, 1998; Ford, 1982; Hawkins, 1971). The second-most likely location is right after the first word of a constituent, particularly when that word is short and highly frequent (e.g., *the*). Clark and Wasow (1998) hypothesized that speakers might commit to producing an utterance by uttering a constituent-initial function word. Doing so would require few processing resources (because function words are easy to retrieve; Bock, 1989; Garrett, 1975; Dell, 1986), and the result would be more time to formulate the material that followed. Of course, word finding difficulties (Schacter et al., 1991) can lead to disfluencies anywhere within an utterance, but the distributional data make clear that they are highly likely to occur at the beginnings of constituents (that is, right at the left edge or immediately after a constituent-initial function word). Moreover, the more complex a constituent, the more likely it is to be articulated with a disfluency (Clark & Wasow, 1998). Thus, a disfluency would be more likely to occur at the beginning of *the red dot* than *the dot*, and even more likely at the beginning of *the red dot that is on the left*.

Given these facts about the distribution of disfluencies in speech, we reasoned that if listeners encountered a disfluency in a location that was a possible (but ambiguous) boundary between constituents (e.g., between two clauses), the parser might opt to build a structure that included a boundary at that location.[2] In order to test this hypothesis, we took advantage of another type of garden path structure, the coordination ambiguity shown in (2). The ambiguity arises because the NP *the waiter* at first appears to be part of a conjoined object, but then turns out to be the subject of a second clause coordinated with the first.

(2)	a. ***Baseline***:	Sandra bumped into the busboy and the waiter told her to be careful.
	b. *Helpful disfluency*:	Sandra bumped into the busboy and the uh uh waiter told her to be careful.
	c. *Unhelpful disfluency*:	Sandra bumped into the uh uh busboy and the waiter told her to be careful.
	d. *Words in the "helpful" location*:	Sandra bumped into the busboy and the tall and friendly waiter told her to be careful.
	e. *Words in the "unhelpful" location*:	Sandra bumped into the busboy and the waiter who was friendly told her to be careful.

Table 2
Percentage of sentences judged grammatical for each of the conditions represented in (2)

Condition	%
Baseline	94
Helpful disfluency	98
Unhelpful disfluency	90
Words in “helpful” location	90
Words in the “unhelpful” location	95

In this experiment, sentences were spoken so that a disfluency occurred either not at all (2a), before the second noun phrase involved in a coordination (e.g., *waiter* in 2b), or before the first noun of the coordination (e.g., *busboy* in 2c). The disfluency before the second NP is hypothesized to be helpful because it is consistent with the two clause structure that the parser needed to settle on once the parse was complete. If disfluencies tend to occur right before or one word into a complex constituent, then the *uh*s in (2b) might serve as a cue to the parser to assume *the waiter* begins a new clause and that the coordination involves two clauses rather than two NPs. A disfluency before the first NP (as in 2c) was predicted to be unhelpful because, again, the *uh*s suggest a complex constituent, but the clause ultimately ends right after the disfluency. As a result, the disfluency is misleading with respect to the ultimately correct analysis. We predicted that sentences with consistent disfluencies would be judged grammatical more often than those with inconsistent disfluencies, because the parser tries to make use of this information when it builds a syntactic structure for a sequence of words.

The results supported these predictions. Sentences in the condition illustrated in (2b) were more likely to be judged grammatical than those illustrated in (2c). Control utterances with modifiers in positions corresponding to the helpful and unhelpful disfluencies (2d and 2e, respectively) did not show this same pattern, which indicates that disfluencies are used as cues, but words in the same locations are not (Table 2).

In more recent work (Bailey & Ferreira, 2003a, 2003b), we have shown that this information is used online and not just to come up with a final interpretation of a sentence. The experiments employ the so-called *visual world paradigm* (Spivey et al., 2002; Tanenhaus et al., 1995; Trueswell, Sekerina, Hill, and Logrip, 1999) in which subjects manipulate objects in response to spoken commands. Participants heard utterances such as either *put the uh uh frog on the napkin in the box*, or *put the frog on the uh uh napkin in the box*. Again, because disfluencies tend to occur before or one word into a complex constituent, the first example should lead participants to look at a frog that is sitting on a napkin, and the second should lead them to expect reference to a napkin that sits inside a box. This is what we observed. In the first case, after people heard the filled pauses and the head noun *frog*, they glanced over to the frog that had to be described with a complex phrase rather than with just a simple NP. It appears, then, that this cue function associated with disfluencies operates early during parsing.

What have we learned from the studies we have discussed thus far? First, disfluencies clearly are not simply filtered before parsing begins. As was mentioned earlier, for a variety of reasons this possibility was never very plausible, but our work demonstrating the effects of disfluencies on parsing makes very clear that the filtering view is untenable. Second, we have learned that disfluencies can prolong the parser's commitment to a given structure, because while the disfluency occurs the structure built up to that point gains strength. Finally, if a disfluency is found in a sentence it is likely to be located right before or one word into a complex constituent, and it appears that the parser uses this information to help it resolve a syntactic ambiguity. But thus far we have considered only filled pauses. How might the parser cope with other kinds of disfluencies, particularly ones that introduce unwanted lexical content?

3.2. Verb substitutions and parsing

Disfluencies provide us with a unique opportunity to investigate the assumption that verb argument structure information is used during parsing, which of course is a fundamental idea in L-TAG (Joshi & Schabes, 1997). That is, a parser which uses information stored in the form of lexically anchored trees would automatically activate the arguments associated with a verb when one is encountered in the input. Normally, these structures would then be available to facilitate parsing, because subsequent input could be integrated into a tree already retrieved based on the verb.

Disfluencies involving a corrected verb provide us with an opportunity to investigate this assumption, and in a unique and original way. Consider again the example *Mary put uh dropped the frog*. After encountering the *uh* and then the word *dropped*, the parser might assume that it has a disfluency on its hands. As described in the earlier section summarizing psycholinguistic investigations of disfluencies, the parser must identify the reparandum and distinguish it from the repair. Assuming it is able to do so (and later we will describe a procedure by which the parser can make this distinction), the parser's next task is to delete the material associated with the reparandum.

But is the reparandum erased in its entirety? Given what we now know about garden-path repair, it is unlikely that the answer to this question is yes. Somewhat surprisingly, it turns out that when people build an incorrect syntactic structure because they have been led down a syntactic garden-path, they often retain the misinterpretation associated with that misanalysis (Christianson, Hollingworth, Halliwell, & Ferreira, 2001; Ferreira, Christianson, & Hollingworth, 2001). For example, given the sentence *While Anna bathed the baby played in the crib*, comprehenders know that the baby played in the crib, but they also tend to believe that Anna bathed the baby (rather than herself). This phenomenon of incomplete reanalysis highlights the fallibility of the psycholinguistic and cognitive operations that perform garden-path repair. Now consider the example we have been discussing, *Mary put uh dropped the frog*. The argument structure of *drop* **can** be compatible with that of *put*, because *drop* allows a theme and a locative prepositional phrase. Therefore, perhaps the already activated argument structure for *put* primes a three-place frame, and this structure would not necessarily be cleanly erased. As a result, the entire sentence might appear to be ungrammatical, because *drop* seems to be missing an argument.

To see how verb corrections are handled during parsing, we conducted a set of experiments using materials exhibiting the main verb/reduced relative (MV/RR) ambiguity (Lau & Ferreira, under review). Sentences beginning with *The little girl selected* ... are ambiguous between a main verb reading like *The little girl selected the right answer, so her teacher gave her a* *prize* or a reduced relative reading like *The little girl selected for the role celebrated with her parents and friends*. Much experimental work has shown that the main verb reading is generally favored by the parser (Bever, 1970; Ferreira & Clifton, 1986; Frazier, 1978; MacDonald, 1994), so if the sentence is disambiguated to the less preferred reduced relative reading, the parser experiences a measurable amount of processing difficulty. In contrast, reduced relatives with verbs that have separate forms for the two structures (e.g., *chose* and *chosen*) do not give rise to this ambiguity, and thus the reduced relative reading causes almost no problems in the processing of a sentence like *The little girl chosen for the role celebrated with her parents and friends*.

Our prediction in these experiments was that even if the unambiguous verb form was present only as a disfluency that was quickly corrected to an ambiguous verb form (*The little girl chosen-uh, selected for the role* ...), it would still exert a disambiguating effect that would lessen the difficulty experienced from the garden path structure. We could compare this to the case where the disfluency verb displays the same ambiguity in form and the same main verb preference, as in *The little girl picked-uh, selected for the role*; in this case we would predict no such effect.

As in most of the experiments on disfluencies that we described above, the task for participants in these two experiments was to listen to sentences one at a time and to press a button to make a judgment of the sentence's acceptability. Again, participants were told not to make their judgments on the basis of the presence or absence of disfluencies. In the first experiment, subjects heard experimental sentences spoken with normal prosody from four conditions. The first two were like the fluent ambiguous and unambiguous examples above, which gave us a baseline for the strength of this type of garden path in spoken materials. The second pair of conditions was disfluent, and these formed the crucial comparison. One condition was made up of sentences like the one shown in (3a). These cases included a verb-replacement disfluency in which the initial, "wrong" verb was an unambiguous form such as *chosen.* In the other condition, the sentences included the same type of disfluency, but in this case the "wrong" verb was another ambiguous form, as in (3b). We predicted that, even though the final, disfluency-excised version of these sentences was the same, the presence of the unambiguous verb form in the disfluency would influence processing and make the garden path in this condition weaker.

(3) a. *Unambiguous disfluency*: The little girl chosen-uh, selected for the role celebrated with her parents and friends.
b. *Ambiguous disfluency*: The little girl picked-uh, selected for the role celebrated with her parents and friends.

In a second, very similar experiment we preserved the methodology and the basic design, but in this case used materials which were disambiguated to the preferred, MV reading, and so were not in fact garden path sentences at all (e.g., *The little girl selected the right answer so her teacher gave her a prize*). However, we hypothesized that the presence of a disfluency including a verb form like *chosen* that was inconsistent with this preferred reading would cause a kind of reverse garden path where getting the usually preferred MV reading would actually cause difficulty. We tested this hypothesis by comparing responses to sentences with this kind of 'inconsistent' disfluency and sentences with a 'consistent' disfluency, as shown below in (4).

(4) a. *Inconsistent disfluency*: The little girl chosen-uh, selected the right answer so the teacher gave her a prize.
b. *Consistent disfluency*: The little girl picked-uh, selected the right answer so the teacher gave her a prize.

The data from these two experiments are presented in Fig. 1.

As Fig. 1 shows, our predictions were supported. In the first experiment, we found a small but significant effect of disfluency type on the proportion of sentences judged grammatical, such that sentences with unambiguous *chosen*-type disfluencies were more likely to be judged grammatical than sentences with ambiguous *picked*-type disfluencies (see Fig. 1a). This mirrored the baseline garden path effect we found between fluent unambiguous and ambiguous controls. In the second experiment, we found an even larger effect of disfluency type, such that sentences with disfluencies including verbs that had an argument structure inconsistent with the final reading were much more likely to be judged ungrammatical than those including "consistent" verb disfluencies (see Fig. 1b). This result was especially interesting because the final reading was actually the preferred MV structure, which was at ceiling levels in both the unambiguous and ambiguous fluent controls.

These findings indicate that the structural information present in word-based disfluencies like substitutions is not completely eradicated during the processing of these disfluencies, and thus exerts an effect on the parse of the sentence. However, from these data it is still not clear whether the effect is mediated by incomplete deletion of an already-built structure into which the corrected verb is simply inserted, or by incomplete deletion of a verb's lexical entry leading to simple priming of verb argument structure (an effect reported by Trueswell & Kim, 1998, in a reading study that used fast priming to make available the two verbs). To try to answer our questions about mechanism, then, we needed a design which disentangles these two possibilities. This brings us to our next experiment.

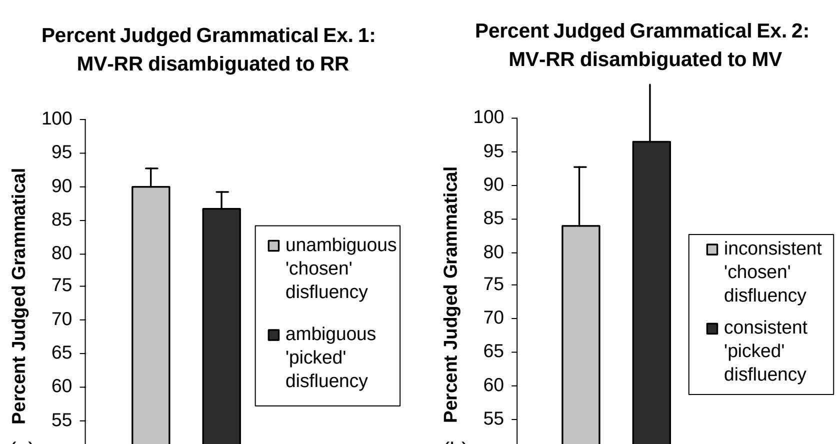

Fig. 1. Percentage of sentences judged grammatical for conditions described in (3) (Experiment 1) and (4) (Experiment 2). Error bars represent standard errors.

3.3. More effects of verb replacements during parsing

In a follow-up study, we again looked at the effect of the verb replacement on parsing, this time using verbs that varied on a different dimension: the number of obligatory arguments associated with them. A 2/3-argument verb such as *drop* can take a third, location argument, as illustrated in the sentence *I dropped the frog in the box*, but the third, goal argument is optional, as we can see from the perfectly grammatical sentence *I dropped the frog*, which leaves the location unspecified. On the other hand, a 3-argument verb such as *put* must take an overt third goal argument, as evidenced by the contrast between the grammatical *I put the frog in the box* and the ungrammatical *I put the frog*.

In this experiment, we were interested in whether differences in the number of arguments required by the verb involved in a verb substitution might affect the degree to which sentences were perceived as grammatical. The comparisons of interest are presented in (5). Specifically, we predicted that if a 3-argument verb was replaced with a 2/3-argument verb, the parser would be more likely to expect the (now optional) goal argument and would thus be more likely to find the sentence ungrammatical in the absence of this argument as compared to a fluent control. Conversely, if a 2/3-argument verb was replaced with a 3-argument verb, the parser might actually have a slightly diminished expectation for a goal argument, and would in its absence be somewhat more likely to find this ungrammatical sentence grammatical.

(5)	a. *Fluent, 2/3-argument verb*:	Simon says you should drop the frog.
	b. *Disfluent, 3-argument→ 2/3-argument verb*:	Simon says you should put-uh drop the frog.
	c. *Fluent, 3-argument verb*:	Simon says you should put the frog.
	d. *Disfluent, 2/3-argument → 3-argument verb*:	Simon says you should drop-uh put the frog.

While the logic of the experiment is similar to the previously described studies on verb substitutions and the MV/RR ambiguity, the materials differ in several important respects. In the MV/RR sentences used above, the structural options associated with the verb affect syntactic and semantic assignment to entities mentioned before the verb is encountered. For example, the issue in the MV/RR ambiguity is whether the first NP should be assigned a theme or an agent role, so that if a verb replacement occurred, the parser might actually have to go back and change its thematic analysis of that constituent. In the sentences used in this experiment, the structural options associated with the verb affect only the postverbal arguments which have not yet been encountered. Therefore, if we still see an effect of the type of verb involved in the disfluency, we can say that the cost is not solely due to having to change thematic role assignments that were already made. In other words, we will have shown that the verb replacements influence not only syntactic reanalysis, but also ordinary, left-to-right structure-building operations.

More importantly, the materials used in the previous study were such that the effects could have been a result either of a priming mechanism that is not necessarily specific to disfluency processing, or to a lingering effect of activated structures, which would be explicitly associated with the parser's attempts to handle disfluencies. In this experiment, we can discriminate between these two possibilities by comparing the fluent and disfluent ungrammatical conditions, in which the 3-argument verb gets only 2 arguments. In the disfluent case, however, the parser initially gets a 2-argument verb which is then replaced by the 3-argument verb. If the mechanism behind the results of the MV/RR verb replacement experiment were driven by priming, we should see no effect of disfluency in the ungrammatical conditions, because an obligatorily 3-argument verb like *put* shouldn't have a 2-argument option to be primed. However, if the verb replacement effect is driven by lingering structure that was built on the basis of the initial verb, we might predict that the ungrammatical sentences involving a 2-argument verb disfluency would be judged grammatical a higher proportion of the time; in this case the initial structure projected is compatible with just two arguments, even though the structure associated with the replacement verb is not.

Participants were 40 Michigan State undergraduates who received course credit for their participation. As their task, participants were asked to make button-press grammaticality judgments to spoken sentences (KGDB), and as in the experiments described above they were instructed not to base their judgments simply on the presence or absence of disfluencies. All sentences were preceded by the frame *Simon says you should* and subjects were deceived into thinking that these sentences had been produced in a previous "Simon says" experiment (and therefore the disfluencies were naturally produced by those speakers).

Sixteen verbs were selected that had an argument structure similar to that of *put*. For eight of these a locative PP argument was obligatory and for the remainder it was optional. We will refer to these as the *locative* items. Another 16 verbs were selected that had an argument structure like *give*'s. Again, eight required a goal argument, and the other eight took it optionally. These will be called the *dative* items. Thirty-two sets of items were constructed from the *put* verbs and 32 from the *give* verbs (as in (5)). Four lists were created such that each subject saw only one stimulus from each set. Thirty-two filler items were also created.

3.4. Results and discussion

Tables 3–5 show the results for the locative and the dative items separately as well as collapsed over verb type. As the examples in (5) show, in all conditions the sentences ultimately occurred with only an object and no PP. Therefore, in the fluent (control) conditions, we expected sentences for which the PP is generally required would be judged grammatical less often than for those sentences which only optionally take the PP. As can be seen in the tables, this expectation was met, although the lack of a PP appears to be a less severe violation for the dative cases than for the locative cases.

Table 3
Percentage of locative items judged grammatical for each of the conditions described in (5)

Sentence fluency	Argument structure of final verb	
	3-argument	2/3-argument
Disfluent	42.0	88.8
Fluent	38.3	94.4

Table 4
Percentage of dative items judged grammatical for each of the conditions described in (5)

Sentence fluency	Argument structure of final verb	
	3-argument	2/3-arguement
Disfluent	66.4	84.1
Fluent	63.5	91.6

The differences that are crucial to our manipulation are those within each argument structure type, where the only difference between conditions is the presence or absence of a disfluency. We see that in both verb types the direction of the difference in the PP-obligatory condition is reversed in the PP-optional condition, and the interaction is statistically significant ($F_{1,39} = 9.54$, $p < .01$). When the final verb is one that optionally takes a PP (like *drop*), the presence of a disfluency involving a verb that obligatorily takes a goal argument results in a decrease in proportion judged grammatical compared to the baseline fluent case. However, when the verbs are reversed such that the disfluency verb optionally takes a PP and the final verb requires one, we find that the presence of the disfluency results in an increase in the proportion judged grammatical compared to the baseline fluent case. We conclude that the disfluency causes the proportion judged grammatical to move in the direction—more or less grammatical—that the sentence would have been if the verb in the disfluency had turned out to be the one intended.

This pattern of results is another piece of evidence that argument structure information in verb-replacement disfluencies can affect parsing, and in a structure different from the MV/RR garden path examined above. More importantly, these results argue against a simple argument structure priming story, and thus lend support to the idea that the effect arises because the structure from the misarticulated verb lingers. If the proportion judged grammatical for the PP-optional case dropped because the 3-argument structure for the *drop* verb was primed by the *put* disfluency, then we would have seen no effect in the PP-obligatory case, where the verb *put* does not have any alternate argument structure options available to be primed. Since we find an effect in this case as well, the mechanism that we postulate for the processing of these repairs needs to somehow allow material from the argument structure of the initial verb to persist in the structural representation even though the lexical anchor of the verb phrase and indeed the clause has presumably been replaced.

Table 5
Percentage judged grammatical, collapsed across item type for each of the conditions described in (5)

Sentence fluency	Argument structure of final verb	
	3-argument	2/3-argument
Disfluent	54.2	86.4
Fluent	50.9	93.0

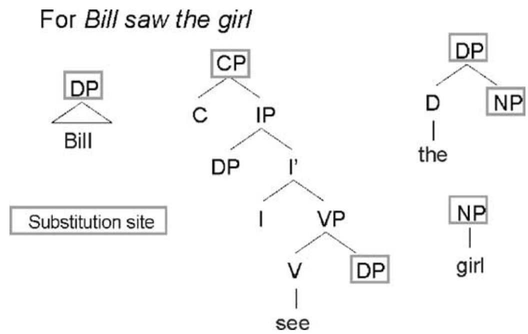

Fig. 2. Basics of human parsing in TAG.

4. A reanalysis model of disfluency processing

In this section, we outline a model of processing that includes some rather simple mechanisms for handling disfluencies. As illustrated in Fig. 2, we will assume the basic representations made available in L-TAG (Joshi & Schabes, 1997).

We assume that words anchor elementary trees, and that elementary trees are combined via Substitution and Adjunction. Substitution is similar to an appending operation: One elementary tree is attached to the bottom node of another elementary tree. The Adjoining operation (which will not be discussed here) essentially inserts a special type of elementary tree inside another elementary tree. We also assume that processing is incremental, which has a very particular meaning in this context (see also Ferreira, 2000): Elementary trees are substituted at the leftmost position possible. This principle means that after a Determiner Phrase is created, for example, it will Substitute into the leftmost position available in the clausal structure made available by the verb, and therefore it will become the clause's subject. In this section we will ignore many important details about parsing, in part because of space limitations, but more importantly to keep the focus on the aspects of the processing that are specifically related to dealing with disfluencies.

Let us begin with the simplest type of disfluency: the unfilled pause, as in *Bill saw the* $\{0\}$ *girl* (the pause is between the determiner and the head noun of the object). Fig. 3 captures how this type of disfluency might interact with parsing operations.

First, the string-initial word *Bill* would be made into a Determiner Phrase (DP). It would then await a structure into which it could be Substituted. That structure becomes available with *saw*, which provides positions for both a subject and an object. Because of incremental processing, *Bill* Substitutes into the subject position. So far, the parser has *Bill saw*, with a slot for the obligatory object. Next, *the* arrives, which brings along its DP with a slot for an NP. Now, based on what we know about the Head Position Effect (Ferreira & Henderson, 1991a, 1998), this DP does not yet attach into the tree; it must wait until the head noun of the NP arrives before Substituting into the clause. At this point we have no choice but to simply make this a stipulation, but it is a necessary one if we are to explain the powerful effect that head position has on parsing (see not just Ferreira & Henderson, 1991a, 1998, but also Christianson et al., 2001; Ferreira et al., 2001; Pickering & Traxler, 1998; Van Dyke & Lewis, 2003). Notice

For *Bill saw the {0} girl*

1. $[\text{Bill}]_{DP}$
2. $[_{IP}\ [\ \__{DP}\ \text{saw}\ \__{DP}]]$
   → parser retrieves the simpler / more frequent argument structure
3. $[_{IP}\ [\ \text{Bill}_{DP}\ \text{saw}\ \__{DP}]]$
4. $[\text{the}\ \__{NP}]_{DP}$
5. {0} → causes delay in substitution; delays commitment to direct object analysis
6. $[\text{the girl}]_{DP}$
7. $[_{IP}\ [\ \text{Bill}_{DP}\ \text{saw the girl}]]$

Fig. 3. Pauses (filled and unfilled) during parsing.

what the pause does: It delays the Substitution of the DP, because it delays receipt of the head noun (compared to an utterance that does not include any material between the determiner and the NP head). When *girl* does arrive, the NP can be Substituted into the DP, and then the DP can be attached into the entire structure. If the sentence did not end there but instead continued with another complement clause verb (as in *Bill saw the uh girl leave* . . .), then the DP *the girl* would only recently have been erroneously placed into the direct object clausal structure, and so this incorrect structure would not have had much time to build up activation and to inhibit other alternatives (including the correct one where *saw* takes a clausal complement).

In this way, we explain the Head Position Effect. In addition, it should be clear why prenominal adjectives and disfluencies such as unfilled pauses and *uh*s behave similarly: They both delay the onset of the head noun and therefore delay (erroneous, in the case of garden-path sentences) attachment, compared to sentences that contain postnominal modifiers. Similarly, if a disfluency occurred **after** the head noun, then attachment would not be delayed (i.e., it would occur right after the determiner was processed, once the head noun was received), and so a commitment to the object analysis would be made sooner. Therefore, if the sentence needed to be reanalyzed (because the DP was not the object but rather the subject of a new clause), then reanalysis would be more difficult due to the prolonged commitment to the erroneous structure. This is what Bailey and Ferreira (2003a, 2003b, in press) observed.

Unfilled pauses and filled pauses (those with *uh*s and *um*s) will behave identically given this mechanism, because neither are lexical items anchoring elementary trees. Notice, by the way, that our approach to the way disfluencies are handled provides a straightforward answer to a rather difficult question that has been asked for about as long as disfluencies in language comprehension and production have been discussed: Are editing terms such as *uh* and *um* words (Clark & Fox Tree, 2002; Fox Tree, 2001)? On our model of disfluency processing, the answer is clearly "no," because our model provides us with a precise definition of what a word (for purposes of the human sentence processing mechanism) is: A word is a bit of linguistic material that anchors an elementary tree. The issue is not that *uh* and *um* fail to take any arguments, because many words have that property. The problem is that *uh* and *um* have no syntactic category, and therefore they lack any sort of representational features that would indicate what slot they should occupy in a phrase-marker.

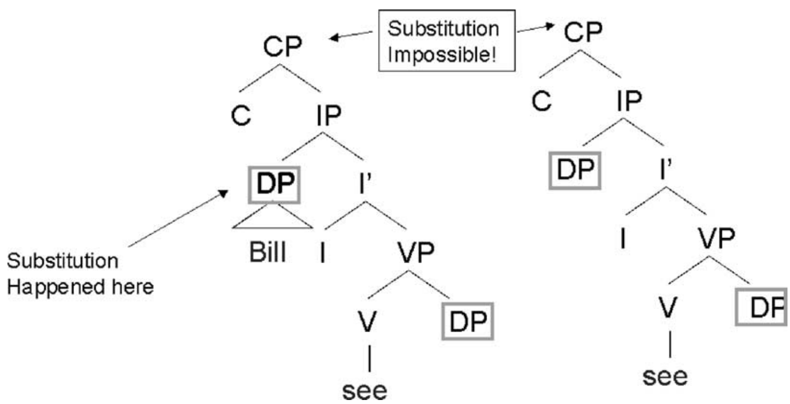

Fig. 4. Two clausal trees that do not allow substitution.

Thus far we have seen that filled and unfilled pauses interact with structure-building operations because they delay the parser's next action; the parser must wait for this kind of disfluency to end and for the next word and tree to become available. As a result, whatever structure the parser has built up to that point gains strength at the expense of any alternatives that could have been pursued instead.

Next, we consider a repeat, as in *Bill saw saw the girl*. Processing proceeds as described above, up to the point at which the second *saw* occurs. We propose that the second *saw* leads to retrieval of an identical elementary tree to the one initially retrieved for *saw*—that is, a clausal phrase marker with a slot for a subject and another for an object. Unfortunately, though, there is nowhere for this second *saw*-headed tree to attach, as shown in Fig. 4.

In this situation, we propose that the parser engages in what we will call **Overlay**. The way Overlay works is as follows: When the parser retrieves any new tree it first looks for a Substitution site. But if no such site is located (which is what will sometimes happen when the utterance contains a repeat or other type of disfluency with lexical content), then the parser looks for node identities and attempts to overlay the trees on top of each other as best it can. This is shown in Fig. 5.

As in the Attach Anyway model of garden-path reanalysis (Fodor & Inoue, 1998, 2000), newer decisions take precedence over older ones. As a result, if the trees differ in any way, it

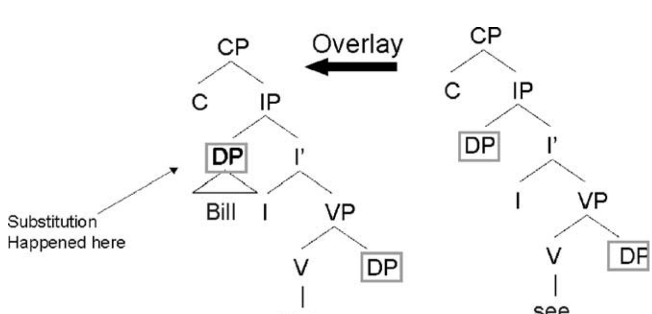

Fig. 5. Overlay: the parser looks for root node identities and, if found, overlays the two trees at root nodes.

is the second, newer tree that will win out (possibly because it has greater activation, having been more recently retrieved), although as the data from the experiment we reported here demonstrate, the other tree is not entirely erased. Indeed, the mechanisms we are proposing here allow us to account for the effect we obtained in the *put–drop* experiment: The reparandum tree has some effect on processing because it was not deleted but rather covered up with the replacement/repair tree. The unique bits of that tree are therefore still somewhat visible to the processor, and so they can affect its operations. Of course, this will not be an issue with the example we are discussing at this point, because the disfluency is a repetition (*saw saw*) and so the two trees that participate in Overlay are identical.

We turn next to consideration of replacements, as in *Bill heard saw the girl*. Replacements are interesting because Overlay will not be a matter of putting identical elementary trees on top of one another. At the very least, the lexical anchor for the interacting trees will be different. Let us begin with such a case: a verb replacement where the two verbs have the same argument structures, as in *Bill heard saw the girl*. Of course, when the parser encounters the second verb, it does not know whether the utterance will turn out to include a disfluency or a coordinated verb (*Bill heard, saw, and spoke to the girl*). Again, for simplicity's sake we will not discuss this possibility here, except to say that: (a) perhaps some sort of Overlay structure IS the way coordinate phrases are built and represented, as argued by Williams (1978; but see Munn, 1993); and (b) frequency might influence the parser's hypothesis about whether it is dealing with a disfluency or a coordination structure, in that simple verb coordination is much less common than DP coordination (for example). Therefore, given a sequence such as *heard saw*, perhaps the parser weights the possibility of a disfluency more heavily because that is consistent with its experience.

Having waved our hands over this problem, let us continue with our discussion of *Bill heard saw the girl*. Processing would proceed as described above for the repetition (*saw saw*) case. What would be different, of course, would be the result: The syntactic forms would overlay perfectly, but the lexical anchor *saw* would sit on top of the lexical anchor *heard*. In other words, the representation would be much like the one shown in Fig. 5, except that the verbal lexical anchors on the two CP trees would be different. For this reason, we might expect that comprehenders would end up more likely to interpret the sentence to mean that Bill saw and possibly also heard the girl than they would if the sentence did not include the replacement. This prediction is one we would like to test in future experiments.

Finally, we consider a disfluency like the one we examined in the experiment we reported in this article: one in which two different verbs are articulated, the first one in error, and where the verbs have different argument structures. As in the case of *saw–saw* and *heard–saw*, the parser would have to engage in Overlay once it discovered that Substitution is impossible. The root CP nodes satisfy the condition on node identity, and so that is the site at which the Overlay will be centered. The result, though, is not perfect overlap, as Fig. 6 illustrates.

This, we believe, is the reason that we get the "lingering" effects of the argument structure from the misarticulated verb. As Fig. 6 shows, that argument structure is not completely covered up by the Overlay operation. As a result, even though *see* does not take a PP argument (on most views of what arguments are; see Koenig, Mauner, & Bienvenue, 2003), it will behave as if it does because the misarticulated verb *put* lurks in the background. This structure will have

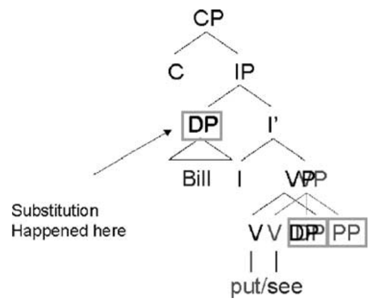

Fig. 6. Overlay for verbs with different argument structures.

a greater or lesser effect depending on how activated it is, which in turn will be a function of how much strength it has gained during structural processing. One of the goals of our future investigations will be to isolate the factors that lead to greater or lesser activation of trees that have been occluded due to the Overlay operation.

One extremely interesting issue that we would also like to pursue concerns the way that garden-path repair and disfluency repair procedures are coordinated. Our initial assumption is that the former take precedence, because garden-path reanalysis operations are much more conservative. Many models of parsing and reanalysis assume that the parser is generally loathe to destroy structure it has already built (e.g., Stevenson, 1998, and references therein), and obviously Overlay is a massively destructive procedure. (In the extreme case, Overlay will result in the parser obliterating all the products of its earlier operations, as in *Mary said Bill—uh Susan told Bill to* . . .) Consider, then, a sentence that contains both a garden-path and a replacement:

(6) Because John coached Bill trained Mary the team did well.

The sequence *because John coached* would be parsed smoothly, but from then on the process is fairly bumpy. The parser will now attach *Bill* as object of *coached*, and of course here it makes its first mistake. The verb *trained* triggers garden-path repair: The parser steals *Bill* from *coached* (Fodor & Inoue, 1998, 2000; Fodor, Ferreira, & Lau, 2003) and constructs *Because John coached, Bill trained*. The parser now awaits an object for *trained*, and it gets one in the form of *Mary*. But now things start to go very badly indeed. At first *the team* seems to be part of some type of coordinate structure (*John coached, Bill trained Mary, the team did drills, and* . . .), but the sentence's ultimate form signals that that assumption is incorrect. Now the parser has a real mess on its hands, and to clean it up it must realize that it has to engage in Overlay: *Bill trained Mary* belongs on top of *John coached*. The final result should be *Because Bill trained Mary, the team did well*.

Of course, it is an empirical question whether the parser could successfully process a sentence like (6), and one we intend to investigate in future work. What is clear, though, is that the processes involved in garden-path recovery and disfluency processing must sometimes work in concert. This point in turn underscores one of the main arguments that we wish to make in this article, which is that our theories of language processing must describe explicitly what the **parser** does when it encounters disfluencies, and that the mechanisms it has available must be transparent with respect to the more "regular" operations that build and reanalyze syntactic structure in fluent utterances. Disfluency processing is not some special problem-solving mode that the human sentence comprehension device goes into; on the contrary, the parser must deal with disfluencies routinely, and its tools for doing so appear to be of the same type as those that are made use of during regular syntactic analysis and reanalysis.

Because of space limitations and also because our research program is just in its infancy, we have left out discussion of a variety of critical and fascinating questions. For example, how can the parser use a filled or unfilled pause as a cue during parsing? How does Overlay work when an *uh* signals that a replacement has occurred? For example, does the parser lower the priority for a coordination analysis when an *uh* separates two words of the same syntactic class (as in *Bill will see uh hear the girl*)? Similarly, would an *uh* or *um* cause the parser to bypass garden-path repair in favor of Overlay when it encounters examples like (6)?

Finally, a significant question that we will focus on in future work is how prosody helps the parser handle disfluencies. It is clear that the parser interprets a disruption to the prosodic integrity of the utterance as a signal that it has a disfluency on its hands (as argued by Levelt & Cutler, 1983). Moreover, it is at least possible that syntactic Overlay involves an analogous process of Intonational Overlay as well. Consider that when a speaker produces a replacement and a correction, the regular intonational contour for the utterance gets disrupted in a highly specific and predictable way: The smooth declination fall (Liberman & Pierrehumbert, 1984) abruptly halts at the interruption site, and the pitch level for the repair is right at about the same place as it was during the reparandum. Therefore, the phonological processor needs to align the intonation in order to create the smooth contour that the speaker intends, and this alignment of the intonational form might help guide the parser to where it should engage in syntactic Overlay. Clearly, this issue is highly complex because it assumes tight communication between the parser and the phonological system, but this idea is hardly controversial (Fodor, 1998; Kjelgaard & Speer, 1999; Steinhauer, Alter, & Friederici, 1999). Indeed, yet another benefit that will come from examining disfluencies carefully is that we stand to learn a great deal about the way that phonological information affects the operations of the parser and the comprehension system as a whole.

5. Conclusions

We have argued that theories of human sentence comprehension must provide a principled account of how people are able to understand sentences with disfluencies, and with such apparent ease. Indeed, the processes are so efficient that often people are not consciously aware that a disfluency even occurred (Lickley & Bard, 1998). Our empirical work has provided some valuable information about the way that disfluencies affect parsing, and we have proposed a theory of disfluency repair which includes a mechanism we term Overlay to explain how revisions and replacements are handled. This research is significant because it not only addresses a highly understudied but significant question about human processing of language, but also because it has clarified some important more general issues. For example, our work on disfluencies and the Head Position Effect suggests that the mechanism underlying the effect is passage of time, which influences activation levels for structural hypotheses. In addition, and more relevant to the theme of this special issue, our research on disfluencies demonstrates once again the power of Tree Adjoining Grammars for virtually all areas of cognitive science. We owe Aravind Joshi a tremendous debt of gratitude for developing this valuable formalism for capturing grammatical knowledge.

Notes

1. Of course, not all silent pauses are disfluencies. We refer here not to silent pauses that might be inserted as part of the normal prosodic form of the sentence (Ferreira, 1993; Selkirk, 1993) but rather those that occur in places that are not licensed by the sentence's prosodic structure, as in the example provided here.
2. This "cuing" explanation of disfluencies also accounts in part for the results we reported in Table 1, because as (1) shows, the head-early disfluency occurs in a predictive location, whereas the head-late disfluency does not. In Bailey and Ferreira (2003a, 2003b) we provide evidence that our results cannot be entirely explained by cuing. We set up an experiment with the same design as (1), but for the conditions represented by (1d) and (1e), disfluencies were replaced with extraneous environmental sounds (which have no cue validity). We found the same pattern of results that was observed with disfluencies, suggesting that the results shown in Table 1 are due *both* to cuing and activation levels associated with the passage of time.

Acknowledgments

This research was supported in part by grants from the National Science Foundation (BCS-9976584, DGE-0114378) and a Michigan State University Distinguished Fellowship awarded to Karl Bailey. The authors thank John Henderson, Timothy Desmet, Susan Brennan, Albert Kim, and an anonymous reviewer for their comments on an earlier draft.

References

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

Arnold, J. E., Fagnano, M., & Tanenhaus, M. K. (2003). Disfluencies signal thee, um, new information. *Journal of Psycholinguistic Research*, *32*(1) 25–36.

Bader, M. (1998). Prosodic influences on reading syntactically ambiguous sentences. In J. Fodor & F. Fereirra (Eds.), *Reanalysis in sentence processing* (pp. 1–46). Dordrecht: Kluwer.

Bailey, K. G. D., & Ferreira, F. (2001). The disfluent hairy dog: Can syntactic parsing be affected by non-word disfluencies? Poster presented at the *14th Annual Meeting of the CUNY Conference on Human Sentence Processing*, Philadelphia, PA.

Bailey, K. G. D., & Ferreira, F. (2003a). Eye movements and the comprehension of disfluent speech. Poster presented at the *16th Annual Meeting of the CUNY Conference on Human Sentence Processing*, Cambridge, MA.

Bailey, K. G. D., & Ferreira, F. (2003b). Disfluencies affect the parsing of garden-path sentences. *Journal of Memory and Language*, *49*, 183–200.

Bailey, K. G. D., & Ferreira, F. (in press). The disfluent hairy dog: Can syntactic parsing be affected by non-word disfluencies? In J. Trueswell & M. K. Tanenhaus (Eds.), *World situated language use: Psycholinguistic, linguistic, and computational perspectives on bridging the product and action traditions*. Cambridge, MA: MIT Press.

Berwick, R., & Weinberg, A. (1984). *Grammatical competence and parsing performance*. Cambridge, MA: MIT Press.

Bever, T. G. (1970). The cognitive basis for linguistic structure. In J. R. Hayes (Ed.), *Cognitive development of language*. New York: Wiley.

Blackmer, E. R., & Mitton, J. L. (1991). Theories of monitoring and the timing of repairs in spontaneous speech. *Cognition, 39*, 173–194.

Bock, J. K. (1989). Closed-class immanence in sentence production. *Cognition, 31*, 163–186.

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.

Bortfeld, H., Leon, S., Bloom, J., Schober, M., & Brennan, S. (2001). Disfluency rates in conversation: Effects of age, relationship, topic, role, and gender. *Language and Speech, 44*, 123–147.

Brédart, S. (1991). Word interruption in self-repairing. *Journal of Psycholinguistic Research, 20*, 123–138.

Brennan, S. E., & Schober, M. F. (2001). How listeners compensate for disfluencies in spontaneous speech. *Journal of Memory and Language, 44*, 274–296.

Brennan, S. E., & Williams, M. (1995). The feeling of another's knowing: Prosody and filled pauses as cues to listeners about the metacognitive states of speakers. *Journal of Memory and Language, 34*, 383–398.

Caplan, D., & Waters, G. (2001). Age, working memory, and on-line syntactic processing in sentence comprehension. *Psychology* & *Aging, 16*, 128–144.

Charniak, E., & Johnson, M. (2001). Edit detection and parsing for transcribed speech. *Proceedings of NAACL*.

Christianson, K., Hollingworth, A., Halliwell, J., & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive Psychology, 42*, 368–407.

Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.

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

Clark, H. H., & Fox Tree, J. E. (2002). Using uh and um in spontaneous speaking. *Cognition, 84*, 73–111.

Clark, H., & Haviland, S. (1977). Comprehension and the given-new contract. In R. Freedle (Ed.), *Discourse production and comprehension* (pp. 1–40). Hillsdale, NJ: Erlbaum.

Clark, H. H., & Wasow, T. (1998). Repeating words in spontaneous speech. *Cognitive Psychology, 37*, 201–242.

Core, M. G., & Schubert, K. (1999). Speech repairs: A parsing perspective. In *Satellite Meeting ICPHS 99* (pp. 47–50).

Cutler, A. (1983). Speakers' conceptions of the function of prosody. In A. Cutler & D. R. Ladd (Eds.), *Prosody: Models and measurement* (pp. 79–91). Heidelberg: Springer-Verlag.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93*, 283–321.

Devescovi, A., Bates, E., d'Amico, S., Hernandez, A., Marangolo, P., Pizzamiglio, L., et al. (1997). An on-line study of grammaticality judgments in normal and aphasic speakers of Italian. In L. Menn (Ed.), *Special issue on cross-linguistic aphasia*. *Aphasiology, 11*(6), 543–579.

Ferreira, F. (1993). The creation of prosody during sentence production. *Psychological Review, 100*, 233–253.

Ferreira, F. (2000). Syntax in language production: An approach using Tree-Adjoining Grammars. In L. Wheeldon (Ed.), *Aspects of language production*. Cambridge, MA: MIT Press.

Ferreira, F., Anes, M. D., & Horine, M. (1996). Exploring the use of prosody during language comprehension using the auditory moving window technique. *Journal of Psycholinguistic Research, 25*, 273–290.

Ferreira, F., Christianson, K., & Hollingworth, A. (2001). Misinterpretations of garden-path sentences: Implications for models of sentence processing and reanalysis. *Journal of Psycholinguistic Research, 30*, 3–20.

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

Ferreira, F., & Henderson, J. (1991a). Recovery from misanalyses of garden-path sentences. *Journal of Memory and Language, 30*, 725–745.

Ferreira, F., & Henderson, J. M. (1991b). The use of verb subcategorization information in syntactic parsing. In G. Simpson (Ed.), *Understanding word and sentence* (pp. 305–330). North-Holland.

Ferreira, F., & Henderson, J. M. (1998). Syntactic reanalysis, thematic processing, and sentence comprehension. In J. D. Fodor & F. Ferreira (Eds.), *Reanalysis in sentence processing* (pp. 73–100). Dordrecht: Kluwer Academic Publishers.

Ferreira, F., Henderson, J. M., Anes, M. D., Weeks, P. A., Jr., & 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.

Fodor, J. D. (1998). Learning to parse? *Journal of Psycholinguistic Research*, *27*(2), 285–319.

Fodor, J. D., Ferreira, F., & Lau, E. (2003). A time-course study of garden path reanalysis. Poster presented at the *16th Annual Meeting of the CUNY Conference on Human Sentence Processing*, Cambridge, MA.

Fodor, J. D., & Inoue, A. (1998). Attach anyway. In J. D. Fodor & F. Ferreira (Eds.), *Reanalysis in sentence processing* (pp. 101–141). Dordrecht: Kluwer Academic Publishers.

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

Ford, M. (1982). Sentence planning units: Implications for the speaker's representation of meaningful relations underlying sentences. In J. Bresnan (Ed.), *The mental representation of grammatical relations* (pp. 798–827). Cambridge, MA: MIT Press.

Fox Tree, J. E. (1995). The effects of false starts and repetitions on the processing of subsequent words in spontaneous speech. *Journal of Memory and Language*, *1995*, 34.

Fox Tree, J. E. (2001). Listeners' uses of um and uh in speech comprehension. *Memory and Cognition*, *29*(2), 320–326.

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

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

Garrett, M. F. (1975). The analysis of sentence production. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 133–175). San Diego, CA: Academic Press.

Godfrey, J. J., Holliman, E. C., & McDaniel. (1992). SWITCHBOARD: Telephone speech corpus for research and development. In *Proceedings of the IEEE Conference on Acoustics, Speech, and Signal Processing* (pp. 517–520). San Francisco, CA: IEEE.

Grossman, M., Mickanin, J., Onishi, K., & Hughes, E. (1996). Verb comprehension deficits in probable Alzheimer's disease. *Brain and Language*, *53*, 369–389.

Hawkins, P. R. (1971). The syntactic location of hesitation pauses. *Language and Speech*, *14*, 277–288.

Heeman, P. A., & Allen, J. (1995). The Trains 93 dialogues. *Trains Technical Note 94-2*, Department of Computer Science, University of Rochester.

Heeman, P., & Allen, J. (1997). Intonational boundaries, speech repairs, and discourse markers: Modeling spoken dialog. In *Proceedings of ACL/EACL*, Madrid.

Henderson, J. M., & Ferreira, F. (2004). *The interface of language, vision, and action: Eye movements and the visual world*. New York: Psychology Press.

Howell, P., & Young, K. (1991). The use of prosody in highlighting alteration in repairs from unrestricted speech. *Quarterly Journal of Experimental Psychology*, *43A*(3), 733–758.

Joshi A. K., & Schabes, Y. (1997). Tree-Adjoining Grammars. In G. Rosenberg & A. Salomaa (Eds.), *Handbook of formal languages* (pp. 69–123). Berlin: Springer.

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

Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000). Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science*, *11*(1), 32–38.

Kjelgaard, M. M., & Speer, S. R. (1999). Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity. *Journal of Memory and Language*, *40*, 153–194.

Koenig, J.-P., Mauner, G., & Bienvenue, B. (2003). Arguments for adjuncts. *Cognition*, *89*, 67–103.

Kowtko, J. C., & Price, P. J. (1989). Data collection and analysis in the airtravel planning domain. In *Proceedings of DARPA Speech and Natural Language Workshop* (pp. 119–125).

Lau, E. F., & Ferreira, F. (submitted) Lingering effects of disfluent material on the comprehension of garden path sentences.

Levelt, W. (1983). Monitoring and self-repair in speech. *Cognition*, *14*, 41–104.

Levelt, W. (1984). Spontaneous self-repairs in speech: Processes and representations. In M. P. R. van den Broecke & A. Cohen (Eds.), *Proceedings of the Tenth International Congress of Phonetic Science* (pp. 105–117). Dordrecht: Foris.

Levelt, W., & Cutler, A. (1983). Prosodic marking in speech repair. *Journal of Semantics*, *2*, 205–217.

Liberman, M., & Pierrehumbert, J. (1984). Intonational invariance under changes in pitch range and length. In M. Aronoff & R. T. Oehrle (Eds.), *Language sound structure: Studies in phonology presented to Morris Halle by his teacher and students* (pp. 157–233). Cambridge, MA: MIT Press.

Lickley, R. J., & Bard, E. G. (1998). When can listeners detect disfluency in spontaneous speech? *Language and Speech*, *41*(2), 203–226.

MADCOW. (1992). Multi-site data collection for a spoken language corpus. In *Fifth DARPA Speech and Natural Language Workshop* (pp. 7–14).

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). Lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*, 676–703.

Marslen-Wilson, W. D., & Tyler, L. K. (1981). Central processes in speech understanding. *Philosophical Transactions of the Royal Society London*, *B295*, 317–332.

McElree, B., & Griffith, T. (1995). Syntactic and thematic processing in sentence comprehension: Evidence for a temporal dissociation. *Journal of Experimental Psychology: Learning Memory* & *Cognition*, *21*, 134–157.

McKelvie, D. (1998). The syntax of disfluency in spontaneous spoken language. *HCRC Research Paper*, HCRC/RP-95.

Mitchell, D. C., & Holmes, V. M. (1985). The role of specific information about the verb in parsing sentences with local structural ambiguity. *Journal of Memory and Language*, *24*, 542–559.

Munn, A. (1993). *Topics in the syntax and semantics of coordinate structures*. Doctoral dissertation, Univeristy of Maryland.

Nakatani, C. H., & Hirschberg, J. (1994). A corpus-based study of repair cues in spontaneous speech. *Journal of the Acoustical Society of America*, *95*(3), 1603–1616.

Oviatt, S. (1995). Predicting spoken disfluencies during human–computer interaction. *Computer Speech and Language*, *9*, 19–35.

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

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

Schafer, A., Carlson, K., Clifton, C., Jr., & Frazier, L. (2000). Focus and the interpretation of pitch accent: Disambiguating embedded questions. *Language and Speech*, *43*, 75–105.

Schacter, S., Christenfeld, N., Ravina, B., & Bilous, F. (1991). Speech disfluency and the structure of knowledge. *Journal of Personality and Social Psychology*, *60*, 362–367.

Selkirk, E. (1993). Sentence prosody: Intonation, stress and phrasing. In J. Goldsmith (Ed.), *Handbook of phonological theory* (pp. 550–569). Basil, Blackwell: Oxford.

Shapiro, L. P., Zurif, E. B., & Grimshaw, J. (1987). Sentence processing and the mental representation of verbs. *Cognition*, *27*, 219–246.

Shriberg, E. E. (1996). Disfluencies in SWITCHBOARD. In *Proceedings of International Conference on Spoken Language Processing*, Addendum (pp. 11–14), Philadelphia, PA.

Shriberg, E. E. (2001). To "errrr" is human: Ecology and acoustics of speech disfluencies. *Journal of the International Phonetic Association*, *31*(1), 153–169.

Smith, V. L., & Clark, H. H. (1993). On the course of answering questions. *Journal of Memory and Language*, *32*, 25–38.

Spivey, M. J., Tanenhaus, M. K., Eberhard, K., & Sedivy, J. C. (2002). Eyemovements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution. *Cognitive Psychology*, *45*, 447–481.

Steinhauer, K., Alter, K., & Friederici, A. D. (1999). Brain responses indicate immediate use of prosody in natural speech processing. *Nature Neuroscience*, *2*, 191–196.

Stevenson, S. (1998). Parsing as incremental restructuring. In J. D. Fodor & F. Ferreira (Eds.), *Reanalysis in sentence processing* (pp. 327–363). Dordrecht: Kluwer Academic Publishers.

Stolcke, A., & Shriberg, E. (1996). Statistical language modeling for speech disfluencies. In *Proceedings of International Conference on Acoustics, Speech and Signal Processing* (pp. 405–409), Atlanta, GA.

Stolcke, A., Shriberg, E., Bates, R., Coccaro, N., Jurafsky, D., Martin, R., Meteer, M., Ries, K., Taylor, P., & Van Ess-Dykema, C. (1998). Dialog act modeling for conversational speech. In *Proceedings of the AAAI-98 Spring Symposium on Applying Machine Learning to Discourse Processing*.

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

Tanenhaus, M., Carlson, G., & Trueswell, J. (1989). The role of thematic structure in interpretation and parsing. *Language and Cognitive Processes*, *4*(3/4), 1211–1234.

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.

Trueswell, J. C., & Kim, A. E. (1998). How to prune a garden-path by nipping it in the bud: Fast-priming of verb argument structures. *Journal of Memory and Language*, *39*, 102–123.

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

Van Dyke, J., & Lewis, R. L. (2003). Distinguishing effects of structure and decay on attachment and repair: A retrieval interference theory of recovery from misanalyzed ambiguities. *Journal of Memory and Language*, *49*, 285–413.

van Wijk, C., & Kempen, G. (1987). A dual system for producing self-repairs in spontaneous speech: Evidence from experimentally elicited corrections. *Cognitive Psychology*, *19*, 403–440.

Williams, E. (1978). Across-the-Board rule application. *Linguistic Inquiry*, *9*, 31–43.