Available online at www.sciencedirect.com
SCIENCE DIRECT®
ACADEMIC PRESS
Journal of Memory and Language 49 (2003) 183–200
Journal of Memory and Language
www.elsevier.com/locate/jml

Disfluencies affect the parsing of garden-path sentences☆

Karl G.D. Bailey and Fernanda Ferreira*

Michigan State University, East Lansing, MI 48824-1117, USA

Received 7 February 2002; revision received 29 January 2003

Abstract

Spontaneous speech differs in several ways from the sentences often studied in psycholinguistics experiments. One important difference is that naturally produced utterances often contain disfluencies. In this study, we examined how the presence of "uh" in a spoken sentence might affect processes that assign syntactic structure (i.e., parsing). Four experiments are reported. In the first, participants judged the grammaticality of sentences that had disfluencies either right before the head noun of the ambiguous phrase or after (e.g., *Sandra bumped into the busboy and the uh uh waiter told her to be careful* or *Sandra bumped into the busboy and the waiter uh uh told her to be careful*). Sentences in the latter condition were judged grammatical less often. This result was replicated in the second experiment, in which disfluencies were replaced with environmental sounds. These findings suggest that interruptions can affect syntactic parsing, and the content of the interruption need not be speechlike. In Experiments 3 and 4 we tested whether these effects occurred because listeners use interruptions as cues to help resolve a structural ambiguity. Results from these latter two grammaticality judgment tasks suggest that when an interruption occurs before an ambiguous noun phrase, comprehenders are more likely to assume that the noun phrase is the subject of a new clause rather than the object of an old one, and furthermore, it appears that the parser is relatively insensitive to the form of the interruption. We conclude that disfluencies can influence the parser by signaling a particular structure; at the same time, for the parser, a disfluency might be any interruption to the flow of speech.
© 2003 Elsevier Science (USA). All rights reserved.

The processes involved in speaking and in writing differ substantially from each other, and so the products of the two systems are not the same. Nevertheless, most psycholinguistic studies of comprehension have used the sorts of linguistic materials that would be produced by a writer rather than a true speaker (Clark, 1997). Although speakers generally aim for an ideal delivery when they speak—that is, they attempt to produce a grammatical, fluent utterance that is similar to what a good writer would produce (Clark & Clark, 1977)—examination of spontaneous speech reveals that the output often falls short of this ideal. One important deviation is the presence of disfluencies in natural speech, which can include long silences, repeated words or groups of words, false starts, abandoned words, and the focus of the present study, items such as "uh" (and its close cousin, "um"; Smith & Clark, 1993). The single utterance below contains examples of almost all of these disfluency types.

(1) But I think uh- uh- uh- precisely because technology itself is certainly more and more... animated... uh and and is... moving faster and faster... uh that's I

☆ This research was supported in part by a Michigan State University Distinguished Fellowship awarded to Karl Bailey and by a grant from the National Science Foundation (BCS-9976584) awarded to Fernanda Ferreira. The authors thank Tom Beckius and Sarah Post for their help with running the experiments and preparing stimuli. Thanks also to Tom Carr, John Henderson, and Alan Munn for their guidance on this project.

* Corresponding author. Fax: 1-517-353-1652.
E-mail address: fernanda@eyelab.msu.edu (F. Ferreira).

0749-596X/03/$ - see front matter © 2003 Elsevier Science (USA). All rights reserved.
doi:10.1016/S0749-596X(03)00027-5

think that's one reason that the concept of the m- the meme uh has some valence these days uh and and and people are are are f- find it intriguing at least (from the National Public Radio program *Talk of the Nation*, May 20, 1999).

Disfluencies are fairly common: One estimate is that they occur at the rate of about five to six per hundred words of spontaneous speech (Bortfeld, Leon, Bloom, Schober, & Brennan, 1999; Dollaghan & Campbell, 1992). Among college professors, rates of disfluencies are higher in fields in which there are more possible ways of describing an idea (Schacter, Christenfeld, Ravina, & Bilous, 1991). In the humanities where there are multiple terms for the same idea, rates of disfluency during lectures are 4.85 per minute. In the physical sciences, on the other hand, where concepts tend to be associated with technical and specific terms, rates of disfluency during lectures are about 1.45 per minute. Given how common disfluencies are in any type of discourse, then, it seems plausible that the language comprehension system must have developed ways to handle them during processing. There are two general views of how disfluencies are processed during comprehension. One assumes that the system somehow filters out disfluencies and then conducts its operations on the sanitized result. The other approach allows for the possibility that listeners actually make use of disfluencies. For instance, based on the presence of items such as "uh," the listener could make inferences (not necessarily consciously) about the speaker's mental state and intentions (Smith & Clark, 1993). We will briefly review these two approaches before describing our own interest in studying disfluencies during comprehension.

There are reasons to think that it might be possible for the comprehension system to filter disfluencies. It has been reported that the prosody of an utterance remains intact after a disfluency is digitally removed (Fox Tree, 1995; Fox Tree, 2001; Levelt, 1984). This result suggests that the speaker somehow sets the disfluency apart prosodically from the rest of the sentence, perhaps leaving cues that the listener can use to locate the disruption's beginning and end points. Indeed, in word gating experiments in which participants were required to state whether a disfluency had occurred, listeners were able to predict that a repair was imminent even before they could identify the first word of the repair; that is, subjects could detect that the repair of a disfluency had been initiated (Lickley & Bard, 1998) before any lexical information was available. Recognizing the beginning of the repair would allow the listener to differentiate between the disfluency and the post-disfluency utterance. In addition, computational linguists have identified several cues that could, in combination, lead to detection of disfluencies by computers, as systems for automated speech recognition also need to handle spoken utterances that deviate from ideal delivery (e.g., Nakatani & Hirschberg, 1994; Shriberg, Bear, & Dowding, 1992).

It has also been shown that people have difficulty remembering where disfluencies in spontaneous speech occurred. When asked to transcribe (Lickley, 1995; Lickley & Bard, 1996) or repeat verbatim (Martin & Strange, 1968) utterances that contained disfluencies, subjects tended either to miss the disfluencies or to report incorrectly that the disfluencies had occurred at clause boundaries. In other words, it is difficult for listeners to consciously identify the location of disfluencies even when their attention is directed towards them. This result is compatible with the suggestion that sentence comprehension mechanisms may filter disfluencies, thereby making them inaccessible to the conscious analysis procedures that support accurate transcription or verbatim repetition. Taken together, the prosodic well-formedness of filtered utterances, the ability of listeners to detect the initiation of a repair, the presence of possible cues for the recognition of disfluencies, and the inability of listeners to consistently locate and identify disfluencies all support the idea that at some stage of processing filtering of disfluencies is possible, at least in principle.

The second approach to disfluencies in comprehension is to assume not that they are filtered, but that they in fact are processed and that they affect the comprehension system. Fox Tree (1995) used a word monitoring task to examine this possibility. Participants were asked to listen to sentences and to simultaneously monitor for the presence of a word specified prior to utterance onset. The experiment suggested that disfluencies in which words are repeated as in (2) do not affect latencies to monitor for the target word *dot*, whereas disfluencies involving a false start as in (3) do slow down monitoring times (Fox Tree, 1995).

(2) Move to the **to the** blue dot on the left (repeat disfluency).

(3) Move to the **green-** blue dot on the left (false start with repair disfluency).

It is possible that processing the false start requires more resources at some level, thus increasing word monitoring times. For example, the effort of holding in working memory the correct material that was uttered right before the false start (i.e., before *green*) could decrease the resources available for the word monitoring task (Fox Tree, 1995). Repeated words, on the other hand, do not mislead the lexical access system and therefore do not interfere with comprehension of the early portion of the sentence.

In addition, listeners' assessments of how much a speaker knows about a topic are influenced by the presence of disfluencies (Brennan & Williams, 1995). For instance, if a speaker is asked *In which sport is the Stanley Cup awarded?* and replies *Uh, hockey* instead of just *Hockey*, the listener is more likely to believe that the speaker is not confident about his answer. On the other hand, if the speaker responds *Uh, I don't know* instead of just *I don't know*, the listener would be more likely to believe that the speaker really did know the answer but was having difficulty retrieving it. Thus, disfluencies affect people's assessments of the speaker's knowledge concerning the topic he or she is talking about.

Disfluencies may also affect comprehension in less metalinguistic ways. If listeners are asked to move to a specific colored dot, a disfluency sometimes makes the task easier (Brennan & Schober, 2001). Specifically, when a command utterance included a false start coupled with an "uh" (as in (4)), participants were faster and more accurate to move to a colored dot, compared to an utterance with just the false start and no "uh."

(4) Move to the purple uh yellow dot.

Even though the addition of an *uh* could, in theory, add to the disruption caused by the disfluency, it seems that the presence of the *uh* actually reduced the effects of removing the incorrect color word contained within the disfluency from the representation of the utterance's meaning. One possible explanation is that the "uh" is an explicit cue that the preceding word was incorrectly spoken and should be ignored; without the "uh," the listener must infer that the word *purple* is wrong and that only *yellow* is to be made part of the ultimate representation for the utterance.

There is evidence to suggest, then, that listeners do not filter disfluencies, but instead process them and even use them as information (both helpful and misleading). In some ways, the absence of filtering is not surprising, because it is difficult to imagine how such a process could be reconciled with incremental interpretation. A great deal of evidence now suggests that as a word is encountered and integrated into a syntactic structure, it is immediately semantically interpreted together with whatever words came before (Altmann & Kamide, 1999; Marslen-Wilson & Welsh, 1978; Sedivy, Tanenhaus, Chambers, & Carlson, 1999; Traxler, Bybee, & Pickering, 1997). A filter, on the other hand, implies a process that buffers an utterance, mentally removes any disfluencies, and only then proceeds with interpretation. Of course, the studies of syntactic parsing cited above used ideal renditions of sentences, and so it is at least in theory possible that incremental interpretation happens only for utterances that do not contain disfluencies (or that a participant expects will not have disfluencies, based on his or her expectations about laboratory stimuli). Fortunately, this implausible idea is inconsistent with the data we will report in this article. It appears, then, that despite the availability of cues that might have been useful to a hypothetical filter, in reality disfluencies are processed by the human comprehension system. In all likelihood, the evidence that people are not good at reporting the presence and location of disfluencies simply reveals that the ultimate product of the comprehension mechanism is a semantic representation for the utterance, not a copy of its surface form.

The critical question, then, is precisely how are disfluencies handled by the language comprehension system? Consider a fairly standard model of comprehension such as the one proposed by MacDonald, Pearlmutter, and Seidenberg (1994), which assumes the existence of a mechanism for accessing information from the lexicon, a syntactic analyzer (henceforth, parser) for building a syntactic representation, and an interpreter for establishing the global meaning of the sentence. Which components can be influenced by disfluencies, and how? In the present work, we focus on the possibility that disfluencies influence a pivotal mechanism in the sentence comprehension system: the parser. If we assume that disfluencies are not filtered prior to processing, then they are part of the input to the syntactic parser. If the parser incorporates words into a syntactic tree incrementally, we can then ask how the parser is affected by the presence of a disfluency before or after a particular word. In order to describe how the parser might deal with disfluencies, we will first summarize what is known about the way the parser processes non-disfluent utterances. Most research on parsing has been conducted with reading paradigms; thus, models of parsing do not currently take disfluencies into account. The general view of parsing is that comprehenders build structure incrementally as they hear or read a sentence. In some situations, the parser encounters problems in assigning structure. When faced with temporarily ambiguous (garden-path) sentences that allow multiple structures to be built, the parser must make a commitment to one of the structures, either by selecting that structure or by assigning that structure greater activation than any alternatives. If the structure that the parser has built is confirmed when the disambiguating word is encountered, parsing will proceed smoothly. If, on the other hand, the structure is incompatible with the disambiguating word, the parser will be forced to reanalyze the material that came before. If reanalysis fails (for whatever reason), subjects will judge the grammatical garden path sentence to be ungrammatical because they cannot find a legitimate parse for it (see Ferreira & Henderson, 1991, for a more detailed description of this logic).

In garden path sentences that begin as in (5), the noun phrase (NP) *the dog* occurs at a point in the sentence where it could be assigned two different roles (i.e., two different structures could be built).

(5) While the boy scratched the dog...

The *dog* could be the object of the first clause, as it is in a sentence such as (6). But in a sentence such as (7), the *dog* is the subject of a second clause. If, in this second case, a comprehender had committed to the direct object structure instead of the less preferred (7), the error would become evident at the disambiguating word *yawned* and the structure would have to be reanalyzed.

(6) While the boy scratched the dog the cat yawned loudly.
(7) While the boy scratched the dog yawned loudly.

The difficulty of reanalysis can be increased with various manipulations of the ambiguous region. If a relative clause occurs after the temporarily ambiguous noun (e.g., *the dog that was hairy*) in a garden path structure such as the one exemplified by (7), the sentences become more difficult to comprehend as measured by a decrease in the number of sentences judged to be grammatical (Ferreira & Henderson, 1991). On the other hand, modifiers (adjectives) placed before the ambiguous noun (e.g., *the hairy dog, the big and hairy dog*) do not significantly affect the difficulty of comprehension. We will refer to this pattern as the head noun position effect.

In this paper, we present evidence that a disfluency that adds no propositional content to a sentence, the filled pause *uh*, can elicit the head noun position effect previously demonstrated with postnominal modifiers. Experiment 1 was an oral analogue of Ferreira and Henderson's (1991) grammaticality judgment task. The head noun position effect was replicated for postnominal disfluencies and modifiers in both types of sentences. The next three experiments were conducted in order to further examine the possible causes of the head noun position effect with disfluencies. First, we examined whether the head noun position effect can be elicited by any interruption, regardless of whether it was speechlike or not. Experiment 2 replaced disfluencies with environmental noises, preserving the interruption, but filling it with a non-speechlike sound. Again, the grammaticality judgment paradigm revealed a head noun position effect, thus indicating that the form of the disfluency in Experiment 1 was not necessarily important, but that the presence of an interruption was. Next, we noted that disfluencies tend to cluster around the beginning of a clause (Clark & Wasow, 1998; Ford, 1982; Hawkins, 1971) and thus could in theory be used as cues to clause boundaries, and those cues would be useful for resolving the syntactic ambiguities studied here. Experiment 3 examined the effect of disfluencies that were either good or bad cues (from the listener's perspective) of upcoming structure in garden path sentences. Disfluencies that were good cues were found to help syntactic parsing relative to bad cues. Modifiers, on the other hand, did not function as cues, suggesting that the head noun position effect found with disfluencies may be wholly unrelated to that obtained with modifiers. Finally, Experiment 4 demonstrated that both disfluencies and environmental noises can cue upcoming structure, again suggesting that the parser may be sensitive to the presence of any sort of interruption. Taken together, the results of the experiments suggest that disfluencies affect the parser because the presence of an interruption predicts the end of a current constituent and the beginning of certain constituent types, and this information may be used by the parser to help it resolve syntactic ambiguity.

Experiment 1

The first experiment was designed to determine whether disfluencies affect the parse of a garden path sentence. In order to examine the processing of garden path sentences, we took advantage of an effect that has been demonstrated with visually presented sentences: the head noun position effect (Ferreira & Henderson, 1991). Recall that modifiers following the head noun of an ambiguous phrase in a garden path sentence increase the likelihood that a sentence will be judged ungrammatical. If disfluencies are part of the input to the syntactic parser, they may also affect the parse of the sentence if the disfluencies occur in the same postnominal position. In addition, this experiment gave us an opportunity to replicate the standard head noun position effect in oral, rather than written language, which would be evidence that the same parsing principles are at work in spoken as in written language comprehension.

Method

Subjects

Thirty undergraduate students from Michigan State University participated in this experiment in exchange for research credit in a psychology class. The subjects were all native speakers of English and were naïve to the purposes of the experiment.

Materials

The stimuli for this and the following experiments were two types of garden path sentences: subordinate-main ambiguities as in (8) and coordination ambiguities as in (9).

(8) While the man hunted the deer ran into the woods.
(9) Sandra bumped into the busboy and the waiter told her to be careful.

For this experiment, only structures ultimately resolved towards the nonpreferred structure were used: For the subordinate-main set, the sentences always contained an intransitive subordinate clause that preceded the main clause, as in (8); for the coordination ambiguity set, the sentences all ended up involving sentence level coordination, as in (9). The subordinate-main ambiguity has been shown to cause a reliable head noun position effect when studied using the grammaticality judgment task in reading (e.g., Ferreira & Henderson, 1991). The head noun position effect has not been studied in the coordination ambiguity structure. Thus, the coordination ambiguity offers a chance to examine whether the head noun position effect is robust across different types of garden-path structures. In addition, some theories of reanalysis suggest that the head noun position effect should not be observed with this structure because the coordination ambiguity is reanalyzed differently than the subordinate-main ambiguity. The reanalysis of this coordination ambiguity is believed to involve the capture of phrasal nodes rather than the theft of terminal elements (nodes which dominate individual lexical items; see Fodor & Inoue, 1998, for details).

In order to test for the presence of a head noun position effect for both modifiers and disfluencies in oral sentence comprehension, we constructed five versions of experimental sentences for subordinate-main structure (see (10)–(14) below; the ambiguous head noun is in bold and the disambiguating word is underlined).[1]

(10) While the man hunted the **deer** <u>ran</u> into the woods (plain ambiguous NP).

(11) While the man hunted the brown and furry **deer** <u>ran</u> into the woods (head noun late; prenominal modifier).

(12) While the man hunted the **deer** that was furry <u>ran</u> into the woods (head noun early; postnominal modifier).

(13) While the man hunted the uh uh **deer** <u>ran</u> into the woods (head noun late; prenominal disfluency).

(14) While the man hunted the **deer** uh uh <u>ran</u> into the woods (head noun early, postnominal disfluency).

The same five versions were constructed for the coordination structure ((15)–(19)).

(15) Sandra bumped into the busboy and the **waiter** <u>told</u> her to be careful (plain ambiguous NP).

(16) Sandra bumped into the busboy and the short and pudgy **waiter** <u>told</u> her to be careful (head noun late; prenominal modifier).

(17) Sandra bumped into the busboy and the **waiter** who was pudgy <u>told</u> her to be careful (head noun early; postnominal modifier).

(18) Sandra bumped into the busboy and the uh uh **waiter** <u>told</u> her to be careful (head noun late; prenominal disfluency).

(19) Sandra bumped into the busboy and the **waiter** uh uh <u>told</u> her to be careful (head noun early, postnominal disfluency).

The stimuli were recorded using Computerized Speech Laboratory (Kay Elemetrics). All utterances were recorded at 10 kHz. The sentences were produced in one intonational phrase (a contour used to control for the effects of prosody described by Kjelgaard & Speer (1999)) in the following way: Each of the experimental sentences was spoken with the adjunct phrase *According to Mary* tacked on to the beginning. This adjunct phrase was followed by a large intonational break, which discouraged a second intonational break from occurring mid-utterance which would have disambiguated the sentence. That is, a large break after *Mary* made it easy for the speaker to produce (for example) *While the man hunted the deer ran into the woods* as a single intonational phrase. As a result, the prosody for the sequence we were interested in (the sentences as in (10)–(19)) had a form more similar to Kjelgaard and Speer's neutral prosody than what they call "cooperating" prosody. (And as our results will show, even if some sort of prosodic break did occur in the middle of the sentences, it was not sufficient to prevent people from being garden-pathed.) The phrase *According to Mary* was subsequently trimmed from the utterance. In order to use the file with the Windows-based experimental software (SuperLab Pro, Cedrus Software), the file was converted to PCM wav format and normalized. For the disfluency conditions, a fairly disruptive form of the filled pause disfluency consisting of two *uh*s was used in order to ensure that the disfluency would not be missed by the subjects (Lickley, 1995).

Phonological analyses of our experimental stimuli revealed that when the nouns were preceded or followed by disfluencies as compared to modifiers, they generally had a slightly longer duration and were spoken at a higher pitch (see Table 1 for characteristics of these stimuli and those used in later experiments). Nevertheless, this pattern did not hold for all conditions. Therefore, we believe the most prudent conclusion is that more work should be done comparing the effects of elicited, natural, and "staged" disfluencies.

Thirty subordinate-main ambiguity sentences and 20 coordination ambiguity sentences were included in this experiment (see Appendix). In addition to the 50 experimental sentences, 50 unambiguous grammatical filler sentences (e.g., *While the scientists shivered it began to snow.*) and 50 ungrammatical filler sentences (e.g., *Eddie hit the ball the man who saw it.*) were included, some of which contained disfluencies. The sentences were pseudorandomized for each subject so that no two experimental sentences would occur in succession. Each subject heard only one version of any particular sentence, but heard each of the five conditions an equal number of times. The experiment was run using SuperLab Pro.

[1] For all items, the prenominal condition modifier condition consisted of two adjectives whereas the postnominal modifier condition was a relative clause with just one adjective. This difference biases the results against the expected effect, which is that the prenominal condition will be easier. In addition, Ferreira and Henderson (1991) found that the postnominal condition was more difficult than the prenominal condition even when the latter was more semantically dense (as it is in these experiments). We set up the experiments in this way in order to avoid the concern that the prenominal condition was easier simply because it contained fewer words.

Table 1
Average duration and F0 for nouns in Experiments 1–4

	Subordinate main ambiguity				Coordination ambiguity					
	Head late, ambiguous (A)		Head early, ambiguous (B)		Head late, unambiguous (C)		Head late, ambiguous (D)		Head early, ambiguous (E)	
	Modifier	Disfl.	Modifier	Disfl.	Modifier	Disfl.	Modifier	Disfl.	Modifier	Disfl.
Duration	**373 ms**	**439 ms**	**350 ms**	**515 ms**	**546 ms**	**434 ms**	**418 ms**	**483 ms**	**418 ms**	**560 ms**
F0	**152 Hz**	**163 Hz**	**158 Hz**	**166 Hz**	**164 Hz**	**152 Hz**	150 Hz	154 Hz	**160 Hz**	**152 Hz**

Positions (A)–(E) refer to the placement of material in the following example sentences: While the man hunted the (A) deer (B) ran into the woods. Sandra bumped into the (C) busboy and the (D) waiter (E) told her to be careful (Significant differences in **bold** (paired *t* tests), $p < .05$.)

Procedure

Before beginning the experiment, subjects were shown examples of grammatical and ungrammatical sentences. Subjects were also told that, just as in spontaneous speech, they should not treat the mere presence of a filled pause as evidence of ungrammaticality. The subjects were then seated in front of a computer monitor and told that they would hear sentences sampled from natural speech. The subjects completed sixteen practice trials (eight grammatical and eight ungrammatical) to insure that they understood the procedure. For each experimental trial, the subjects pressed a button in order to hear the sentence. Once the sentence was complete, the subjects judged the grammaticality of the sentence by either pressing a button labeled "Grammatical" or a button labeled "Ungrammatical." Judgments were automatically recorded by the experimental software.

Design and analysis

The subordinate-main and the coordination structures were analyzed separately. For each structure, a 2×2 ANOVA with head noun position (early or late) as the first variable and intervening material (disfluency or modifiers) as the second variable was conducted. In order to compare the baseline condition to the other conditions, we conducted a one way ANOVA with condition as the variable. The proportion of sentences judged grammatical was the dependent measure in this experiment. Because proportions exceeded 80% in some conditions, the analyses were conducted on arc sin transformed scores (Winer, 1971).

Results and discussion

The results of the modifier conditions for both structures (seen in Figs. 1 and 2) replicated the head noun position effect described in Ferreira and Henderson (1991). The plain ambiguous noun phrase in the subordinate-main structure (the horizontal line in Fig. 1) was judged grammatical 83% of the time. This provides a baseline for the comparison of the prenominal and postnominal conditions. The prenominal modifiers were

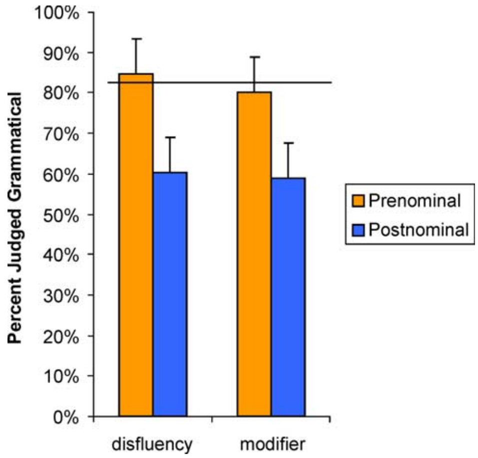

Fig. 1. Percentage of sentences judged grammatical for the subordinate-main structure in Experiment 1. The horizontal black line illustrates the baseline plain ambiguous noun phrase condition. Error bars represent standard errors.

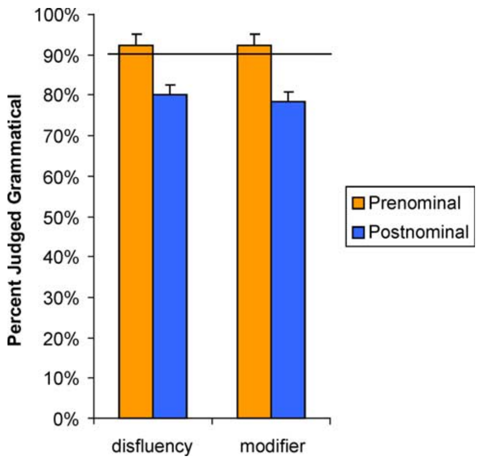

Fig. 2. Percentage of sentences judged grammatical for the coordination ambiguity structure in Experiment 1. The horizontal black line illustrates the baseline plain ambiguous noun phrase condition. Error bars represent standard errors.

judged grammatical 80% of the time. This matches the pattern of data described by Ferreira and Henderson (1991) where the prenominal modifier condition was judged grammatical about as often the plain ambiguous noun phrase condition. The postnominal modifiers were judged grammatical significantly less often (59% judged grammatical) than the prenominal modifiers (by subjects $F_{1,29} = 20.37, p < .01$; by items $F_{1,29} = 13.73$, $p < .01$). The same pattern of data was seen in the modifier conditions for the coordination ambiguity structure (Fig. 2). The prenominal modifier condition (93% judged grammatical) and the plain ambiguous noun phrase condition (90% judged grammatical) were similar, while the sentences in the postnominal modifier condition (78% judged grammatical) were judged grammatical significantly less often (by subjects $F_{1,29} = 9.14$, $p < .01$; by items $F_{1,19} = 8.59$, $p < .01$). These results indicate that the head noun position effect described in garden-path sentences with postnominal modifiers occurs in speech as well as in reading paradigms, and that the coordination ambiguity structure, which some theories state is reanalyzed by a different method than the subordinate-main structure, also shows a head noun position effect during syntactic reanalysis (contrary to the predictions of those same theories; e.g., Fodor & Inoue, 1998).

In addition to replicating the head noun position effect in speech and demonstrating that the head noun position occurs in another structure, Experiment 1 also made our most critical point: that the disfluency *uh* can cause a head noun position effect. The grammaticality judgments in the disfluency conditions showed the same pattern as in the modifier conditions. In the subordinate-main structure (Fig. 1), significantly fewer sentences were judged to be grammatical in the postnominal disfluency condition (60% judged grammatical) than in the prenominal disfluency condition (85% judged grammatical; by subjects $F_{1,29} = 27.41$, $p < .01$; by items $F_{1,29} = 21.16$, $p < .01$). The prenominal disfluency condition was similar to the plain ambiguous noun phrase condition (recall, 83% judged grammatical). The results for the coordination ambiguity structure (Fig. 2) parallel the results of the subordinate-main structure, with significantly more sentences being judged grammatical (by subjects $F_{1,29} = 8.92$, $p < .01$; by items $F_{1,19} = 8.37$, $p < .01$) in the prenominal condition (93% judged grammatical) than in the postnominal condition (80% judged grammatical). Once again, the prenominal condition was similar to the plain ambiguous noun phrase condition (90% judged grammatical). The similarity between the disfluency and modifier conditions is borne out by the presence of a main effect of position (prenominal vs. postnominal) and a lack of a main effect of material (disfluency vs. modifiers) or an interaction between the two variables. The main effect of position was highly significant for both structures (subordinate-main: by subjects $F_{1,29} = 45.99$, $p < .01$; $F_{1,29} = 24.97$, $p < .01$; coordination ambiguity: by subjects $F_{1,29} = 15.17$, $p < .01$; by items $F_{1,19} = 14.72$, $p < .01$).

Experiment 1 shows that parsing of a garden-path sentence is made more difficult when material intervenes between the head noun of the ambiguous phrase and the disambiguating word. At this point it seems that it does not matter whether the material consists of words (modifiers) or disfluencies—both increase the likelihood that a sentence will be judged ungrammatical. It also appears that this effect of head location relative to the disambiguating word is the same in the two different types of garden-path sentences examined here: the subordinate-main structure, and the coordination ambiguity structure. Finally, it is reassuring to know that effects obtained with written materials replicate so directly with spoken utterances. It appears, then, that the principles that influence the interpretation of garden-path sentences and that have been demonstrated in reading studies generalize to spoken language as well.

The critical conclusion is that disfluencies can affect the syntactic parse of a sentence and that they can have much the same effect that words do. One might be immediately tempted to assume that same effects are created by identical causes. Unfortunately, this inference is not valid. We have only demonstrated that modifiers and disfluencies yield the same pattern of results, but this similarity could be a coincidence: Multiple mechanisms could lead to people making the same decision concerning the sentences' grammaticality. In order to establish that the same mechanism produces this pattern for both modifiers and disfluencies, further investigation is required. We can think of at least two fundamentally different ways that disfluencies could affect the syntactic parse of a sentence. First, the effect may somehow be related to the explanations proposed for modifiers (e.g., Ferreira & Henderson, 1998; Fodor & Inoue, 1998). For example, both disfluencies and modifiers delay the onset of the disambiguating word in the postnominal condition syntactic structure relative to the other conditions, as has been observed by Ferreira and Henderson (1998). As a result, the wrong analysis of the noun phrase persists for a longer time, and the persistence of this incorrect parse could cause it to gain strength and could cause the correct parse to lose activation. (Rather than the incorrect parse gaining strength, it could be that the correct, intransitive structure loses activation over time, and, of course, both might be true: The longer the incorrect structure persists, the more strength it acquires and the more it inhibits the ultimately correct analysis. For the purposes of this example, all of these descriptions are equivalent.) In addition, the likelihood of correctly reinterpreting the meaning of a garden-path sentence decreases the longer the wrong parse persists (Christianson, Hollingworth, Haliwell, & Ferreira, 2001). Thus, both the postnominal disfluencies and postnominal modifiers conditions result in a lower proportion of grammatical judgments. Of course, other theories might have other explanations for the head noun position effect; the point that we are trying to make is that the mechanism that causes the head noun position effect in both disfluency and modifier conditions might be the same. Such a hypothesis could be advanced because the disfluencies and the modifiers share one thing in common: They are produced as part of a speech stream by a speaker. Thus, the parser might try to deal with all speechlike sounds in the same way. We will refer to this first possibility as the speechlike hypothesis: The parser deals with all speechlike sounds in the same way and thus disfluencies and modifiers affect the parser uniformly. This hypothesis could be realized as the particular mechanism (delay) that we described earlier, but other explanations (e.g., Fodor & Inoue, 1998) are also possible.

The second possible mechanism to explain the effect of disfluencies on the parsing of garden-path sentences is specific to disfluencies and appeals to the parser's potential use of any type of cue in the speech stream (including disfluencies or interruptions) to predict structure. Several studies have shown that speakers tend to be disfluent either right before they begin a clause, or after a clause's first word (which is often a determiner such as *the*) (Clark & Wasow, 1998; Ford, 1982; Hawkins, 1971). Other studies have shown that speakers tend to displace disfluencies to clause boundaries when asked to transcribe (Lickley, 1995) or repeat (Martin & Strange, 1968) utterances. If the parser is able to use information about the co-occurrence of disfluencies and clause boundaries, then the presence of a disfluency before a noun phrase which could either be a subject of a new clause or an object of an old one could cause the parser to weigh the former analysis more heavily. We will refer to this second mechanism as the signaling hypothesis: Listeners can make use of disfluencies as signals to guide their syntactic parse. (Note that here we make reference to signals purely from the point of view of the listener. It is our contention that the listener may interpret the presence of a disfluency as a signal; we are completely agnostic as to the speaker's intent.)

At this point, we cannot distinguish between the speechlike and the signaling hypotheses, because the results of Experiment 1 are consistent with both. The disfluencies in the prenominal position signal that the noun phrase is a subject (Clark & Wasow, 1998), which is the correct analysis. Therefore, sentences would be judged grammatical more often, because the cue and the structure are consistent. The disfluencies in the postnominal condition signal the start of a new clause and therefore imply that the preceding determiner plus noun sequence was an object. This analysis is incorrect, and so sentences in the postnominal condition would be judged ungrammatical more often because the cue and the structure are inconsistent. Of course, the same explanations would not work for the modifier conditions, but this would simply mean that the similar-looking effects for disfluencies and modifiers actually are created by different sentence processing mechanisms. This is a major difference between the signaling and speechlike hypotheses. However, in all cases the disfluencies occur in a position where appealing to a single mechanism for disfluencies and modifiers (and thus rejecting the signaling hypothesis) could also explain the pattern of results.

We ran three additional experiments to determine whether the speechlike hypothesis or the signaling hypothesis can explain the effect of disfluencies on the parsing of garden-path sentences. Experiment 2, which we describe next, examines a major assumption of the speechlike hypothesis: that interruptions must be sounds the speaker creates.

Experiment 2

The speechlike hypothesis assumes that material following the head noun of an ambiguous noun phrase makes parsing difficult for the same reason whether the material is a disfluency or a modifier, as long as that sound is one associated with speech. One way to test this hypothesis is to include material that occurs immediately following the ambiguous head noun but was clearly not produced by the speaker or by any speaker. We therefore chose to use some unconventional stimuli. In these utterances, disfluencies were replaced with noises that commonly occur in the environment, such as the sounds of dogs barking and doorbells ringing. These environmental sounds were placed in either the prenominal or postnominal position (interrupting the utterance), exactly the same places that had contained words or disfluencies in Experiment 1. Because speakers do not themselves create the environmental noises and have no control over their timing or occurrence, the parser should have no reason to attempt to incorporate them into the syntactic structure. In other words, they should not be used by the parser as information.

It could be argued that these stimuli are not ecologically valid; however, it is clear that many of the complex, written stimuli in psycholinguistic experiments also fall outside the scope of a subject's daily experience with language. Because spontaneous speech often occurs under noisy conditions, listeners will often have the experience of listening to an utterance that stops momentarily because of some auditory intrusion. Certainly, the interruptions in this experiment are somewhat different from many situations in which a very loud noise (in this experiment, the environmental noises were much louder than the preceding and following speech stream) interrupts a speaker, because in this experiment the speaker and the noise do not overlap. In this experiment, however, we are concerned particularly with the effects of any interruption, and we wanted to be certain that subjects could hear all of the words in the sentence (given that they were performing a grammaticality judgment task). Therefore, words and noises were kept distinct.

Method

Subjects

Thirty five undergraduate students from Michigan State University participated in this experiment in exchange for research credit in a psychology class. The subjects were all native speakers of English and were naïve to the purposes of the experiment. These subjects had not participated in any other experiments reported in this paper.

Materials

The stimuli for these experiments were identical to the stimuli in Experiment 2, except that all of the disfluencies were replaced with environmental noises (cats meowing, dogs barking, telephones ringing, doorbells ringing, sneezing, coughing). Excising or replacing disfluencies in a sentence often does not affect its prosody (Clark, 1996; Fox Tree, 1995; Fox Tree, 2001; Levelt, 1984); thus, subjects did not report that the stimuli were unusual (except, of course, for the presence of the environmental noise). The sounds inserted into the utterances were much louder than the preceding or following text, thus making it implausible for the speaker to continue speaking over the sound. Table 2 lists the differences in amplitude (measured in decibels) between the noises, modifiers, and disfluencies and the corresponding head noun. Overall, the modifiers are much louder than the head noun. The noises occurring before a head noun in coordination sentences most likely showed less of a difference because some of these sentences were from Experiment 4. The noises were identical to the noises in other experiments and locations, but they occurred earlier in the sentence, where the surrounding speech was louder.

Thirty subordinate-main ambiguity sentences and 20 coordination ambiguity sentences were included in this experiment (see Appendix). In addition to the 50 experimental sentences, 50 unambiguous grammatical filler sentences and 50 ungrammatical filler sentences were included. The filler sentences were identical to those in Experiment 1, except that the disfluencies in the filler sentences were also replaced with environmental noises. The sentences were pseudorandomized for each subject so that no two experimental sentences would occur in succession. Each subject heard only one version of any particular sentence, but heard each of the five conditions an equal number of times. The experiment was run using SuperLab Pro.

Procedure

The procedure for Experiment 2 was identical to Experiment 1, with the addition of an instruction indicating that the sentences in the experiment included some sentences in which the speaker was interrupted by an environmental noise. Both grammatical and ungrammatical examples that contained an environmental noise were given, and both grammatical and ungrammatical practice sentences contained environmental noises.

Design and analysis

As in the previous experiment, the two structures were analyzed separately. For each structure, a 2×2 ANOVA with head noun position (early or late) as the first variable and intervening material (noises or modifiers) as the second variable was conducted. In order to compare the baseline condition to the other conditions, we conducted a one way ANOVA with condition as the variable. The dependent measure in this experiment was the proportion of utterances judged grammatical. Because proportions exceeded 80% in some conditions, the analyses were conducted on arc sin transformed scores (Winer, 1971).

Results and discussion

Once again, we obtained the head noun position effect for modifiers. In the subordinate-main structure (Fig. 3), the plain ambiguous noun phrase condition (81% judged grammatical, indicated by the horizontal line in Fig. 3) and the prenominal modifier condition (73% judged grammatical) were significantly more likely to be judged

Table 2
Mean amplitude difference (in decibels) between disfluencies, modifiers, and noises and corresponding head nouns in Experiments 1–4

	Subordinate main ambiguity		Coordination ambiguity	
	Head late (dB)	Head early (dB)	Head late (dB)	Head early (dB)
Disfluency	6.97	3.88	3.66	4.31
Modifier	2.11	0.14	1.56	−0.53
Noise	4.52	3.66	1.47	4.73

Positive values indicate that the disfluency, modifier, or noise is louder than the noun.

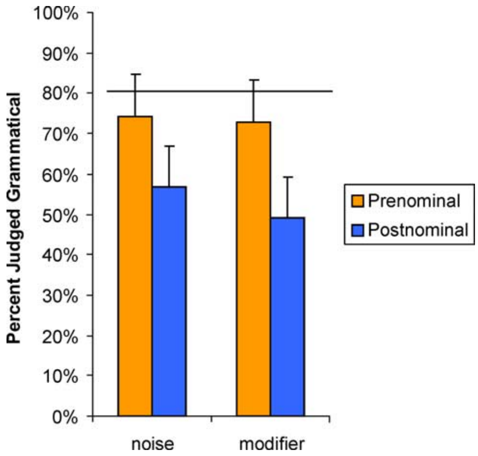

Fig. 3. Percentage of sentences judged grammatical for the subordinate-main structure in Experiment 2. The horizontal black line represents the baseline plain ambiguous noun phrase condition. Error bars represent standard errors.

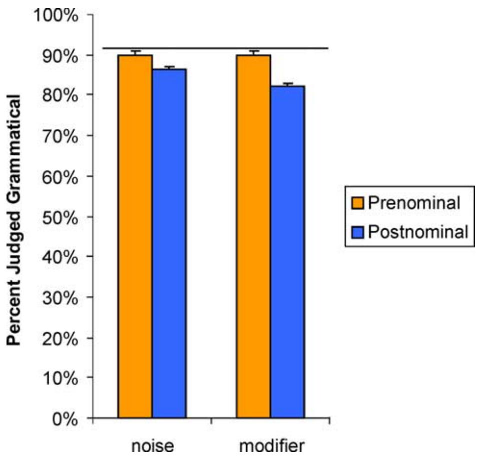

Fig. 4. Percentage of sentences judged grammatical for the coordination ambiguity structure in Experiment 2. The horizontal black line represents the baseline plain ambiguous noun phrase condition. Error bars represent standard errors.

grammatical (by subjects $F_{1,34} = 26.27,\ p < .01$; by items $F_{1,29} = 12.76,\ p < .01$) than the postnominal modifier condition (49% judged grammatical). In the coordination ambiguity structure (Fig. 4), the plain ambiguous noun phrase (91% judged grammatical, indicated by the shorizontal line in Fig. 4) and the prenominal modifier (90% judged grammatical) conditions were judged grammatical marginally more often (by subjects $F_{1,34} = 2.87,\ p = .1$; by items $F_{1,19} = 2.90,\ p = .1$) than the postnominal modifier condition (82% judged grammatical). The results for the coordination ambiguity perhaps failed to reach significance because this structure is easier to parse and reanalyze than the subordinate-main structure, and performance was therefore close to ceiling. However, the pattern of results was the same as we observed in Experiment 1.

Even though the environmental noises were not related to speech in any way, they also elicited a head noun position effect. Sentences with prenominal noises (74% judged grammatical) in the subordinate-main structure were judged grammatical more often (by subjects $F_{1,34} = 12.96,\ p < .01$; by items $F_{1,29} = 16.59,\ p < .01$) than postnominal noises (57% judged grammatical). The same pattern was seen in the coordination ambiguity structure, where the prenominal noise condition (90% judged grammatical) received a higher proportion (although not significant by either items ($F_{1,19} = 0.75,\ p > .1$) or by subjects ($F_{1,34} = 1.51,\ p > .1$)) of grammatical judgments than the postnominal noise condition (86% judged grammatical). Thus, there was an overall main effect of position for the subordinate-main structure (by subjects $F_{1,34} = 45.89,\ p < .01$; $F_{1,29} = 27.28,\ p < .01$) and a marginal main effect of position for the coordination ambiguity (by subjects $F_{1,34} = 3.44,\ p = .07$; not significant by items $F_{1,19} = 2.72,\ p > .1$). Once again, there was no evidence of a main effect of material or an interaction between position and material, implying that noises and words elicited the same pattern of grammaticality judgments.

The results of this experiment suggest that the speechlike hypothesis cannot account for the occurrence of the head noun position effect found with disfluencies. When environmental noises intervene between the head noun of the ambiguous noun phrase and the disambiguating verb, the head noun position effect still occurs, even though the interruption consists of material not under the control of the speaker (that is, few speakers meow or bark during spontaneous speech, and even fewer ring like a telephone; moreover, speakers cannot induce other creatures to make these noises for them). Because the noises are not controlled by the speaker, the listener should not attempt to incorporate them into syntactic structure. Moreover, these noises, while frequent in the environment in which listeners acquire and use language, occur in random sentential locations. Therefore, the noises are not predictive and so should be filtered from the input to the parser. What was found instead was that the environmental noises **did** affect the syntactic parse. Experiment 2 suggests, then, that the head noun position effect found with disfluencies may not necessarily be due to the same mechanism that causes the head noun position effect with modifiers, because environmental noises, which have very little in common with modifiers, elicit the same effect. What environmental noises do have in common with modifiers is the insertion of delay between the ambiguous head noun and disambiguating verb. While this could account for the similar pattern obtained for noises, disfluencies, and modifiers, disfluencies and noises have a particular feature that modifiers do not: they interrupt the stream of linguistic items that need to be parsed. Thus, even though we have demonstrated that the head noun position effect can be elicited by an environmental noise, it remains to be seen whether interruptions and modifiers will pattern differently in a situation where only cueing can affect the syntactic parse. Experiment 3 was designed to directly examine any possible cueing effects of disfluencies in speech.

Experiment 3

To examine whether the position of a disfluency in an utterance can help a listener to disambiguate an upcoming ambiguous structure, it is necessary to place disfluencies in positions that are either consistent or inconsistent with the upcoming syntax of the sentence. Recall that disfluencies are frequently associated with the initiation of a major constituent such as a clause or a complex noun phrase and the closure of the previous constituent. If the parser encountered a disfluency at a point were the current constituent could be closed and a new major constituent initiated, the disfluency might bias the parser in favor of just such an interpretation if the parser could make use of the association between disfluencies and complex structure.

At the same time, it is important that this manipulation of informative versus misleading disfluency location not be confounded with the manipulation used previously, namely prenominal versus postnominal location—otherwise the speechlike and signaling hypotheses cannot be cleanly separated. Recall that the postnominal position was responsible for the decrement in judgment of grammaticality. In this position, material intervenes between the head noun and the disambiguating verb. In order to test cueing (and avoid a confound with a possible mechanism other than cueing that might be at work in the case of modifiers), the distance between ambiguous noun and disambiguating verb must be held constant in all conditions.

It turns out that the coordinate ambiguity structure allows these two influences to be disentangled because consistent and inconsistent disfluencies can be introduced without adding any additional material between the head of the ambiguous noun phrase and the disambiguating word. When the coordination ambiguity is resolved in favor of the non-preferred structure, a disfluency before the second noun phrase in the coordination, as in (20), is a good signal of the closure of a major constituent (in this case a clause) and the initiation of a new major constituent, based on the assumption that the language comprehension system, and specifically, the parser, can make use of the co-occurrence between disfluencies and the start of major constituents. However, if a disfluency before the first NP, as in (21), is to signal any major constituent boundary at all, it must be a conjoined object noun phrase (i.e., the parser's preferred structure, but the one that is ultimately incorrect). The disambiguating word eventually rules out this analysis. Therefore, we would expect a greater proportion of sentences with the disfluencies located as shown in (20) to be judged grammatical in comparison to sentences with the disfluencies located as in (21). This prediction is based on the assumption that a disfluency in the former location serves as a helpful cue to the eventual syntactic structure, which in turn makes the sentence easier to parse and results in an increase in the likelihood of its being labeled as grammatical.

(20) Sandra bumped into the busboy and the uh uh waiter told her to be careful (good signal disfluency).

(21) Sandra bumped into the uh uh busboy and the waiter told her to be careful (bad signal disfluency).

Note that in both cases, the disfluency does not occur postnominally and no material is introduced between the ambiguous noun phrase and the disambiguating word. Therefore, the good signal condition, which is equivalent to the prenominal condition of Experiments 1 and 2, should yield a pattern of results similar to the results from that condition in previous experiments. The bad signal condition has no equivalent in the previous experiments as it does not involve manipulation of the ambiguous NP, but it may lead to an increase in the proportion of sentences judged ungrammatical because it may bias the parser in favor of the ultimately incorrect structure. The disfluency in the "bad signal" condition is "bad" only in the sense that it could serve as a misleading cue to the ultimately correct syntactic structure. It does not, however, lead to an improbable utterance, as word finding difficulties may cause a disfluency to occur before any content word. Of course, if the placement of disfluencies in this experiment has no cueing effect on the parse of the sentence there should be no difference between the two conditions.

Experiment 3 included control conditions comparable to (20) and (21) in which words replaced the disfluencies. Because the signaling hypothesis is concerned with the pattern of co-occurrence between disfluencies and structures, we did not expect prenominal modifiers placed prior to the second noun phrase as in (22) to result in more sentences being judged grammatical than if the prenominal modifier were placed prior to the first noun phrase as in (23). There is no co-occurrence pattern of modification of NPs in a coordinate structure that would give any indication as to the upcoming structure.

(22) Sandra bumped into the busboy and the short and pudgy waiter told her to be careful (good signal equivalent modifiers).

(23) Sandra bumped into the short and pudgy busboy and the waiter told her to be careful (bad signal equivalent modifiers).

Finally, as in the previous two experiments, we included a plain ambiguous noun phrase condition without disfluencies or modifiers (identical to (15)).

Method

Subjects

Thirty undergraduate students from Michigan State University participated in this experiment in exchange for research credit in a psychology class. The subjects were all native speakers of English and were naïve to the purposes of the experiment. These subjects had not participated in any other experiments reported in this paper.

Materials

Twenty coordination ambiguity sentences were included in this experiment (see Appendix). In addition to the four conditions listed above, a baseline condition as shown in (15) was included (no extra words or disfluencies). Twenty unambiguous grammatical filler sentences and 20 ungrammatical filler sentences were included. The filler sentences were selected from those used in Experiment 1. The sentences were pseudorandomized for each subject so that no two experimental sentences would occur in succession. Each subject heard only one version of any particular sentence, but heard each of the five conditions an equal number of times. The experiment was run using SuperLab Pro.

Procedure

The procedure for this experiment was identical to that of Experiment 1.

Design and Analysis

A 2×2 ANOVA with position (good signal position or bad signal position) as the first variable and intervening material (disfluency or modifiers) as the second variable was performed. In order to compare the baseline condition to the other conditions, we conducted a one way ANOVA with condition as the variable. In addition to these ANOVAs, *t* tests were used to compare the two disfluency conditions and the two modifier conditions. The proportion of grammatical judgments was the dependent measure in this experiment. Because proportions exceeded 80% in some conditions, the analyses were conducted on arc sin transformed scores (Winer, 1971).

Results and discussion

Because the coordination ambiguity structure does not produce as strong a garden path as does the subordinate-main structure, mild ceiling effects occurred (all conditions were judged grammatical over 90% of the time). The plain ambiguous noun phrase condition which serves as a baseline (the horizontal line in Fig. 5) was judged grammatical 94% of the time. The ANOVA showed no significant main effects. However, a significant interaction between the disfluencies versus modifiers variable and the position variable (Fig. 5; by subjects $F_{1,29} = 9.81,\ p < .01$; by items $F_{1,19} = 7.63,\ p < .02$) did occur. When the modifier and disfluency conditions are examined separately, the reason for this interaction becomes clear. The bad signal disfluency condition (90% judged grammatical) was judged grammatical significantly less often (by subjects $t_{1,29} = 2.52,\ p < .02$; by items $t_{1,19} = 2.89,\ p < .01$) than the good signal disfluency condition (98% judged grammatical). The two modifier control conditions did not differ statistically from one another (90% in the good signal control condition and 95% in the bad signal control condition; by subjects $t_{1,29} = 1.53,\ p > .1$; by items $t_{1,29} = 1.22,\ p > .1$).

These results match the predictions made by the cueing hypothesis. Modifiers, which have no pattern of co-occurrence in spontaneous speech that could be used to guide the parse of a sentence, show no difference between the good signal equivalent and bad signal equivalent conditions. If there is any effect at all in the modifier controls, it is in the opposite direction to the disfluency conditions. Disfluencies, on the other hand, behave differently. The consistent disfluency condition resulted in a higher proportion of judgments of grammaticality, which would be expected if the syntactic parser was using the occurrence of a disfluency as a signal to a clause boundary. Because that structure is the one that is ultimately correct, it is easier for the parser

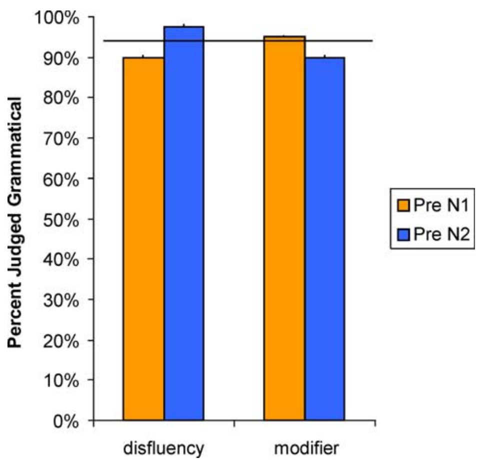

Fig. 5. Percentage of sentences judged grammatical for the coordination ambiguity structure in Experiment 3. The horizontal black line represents the baseline plain ambiguous noun phrase condition. Error bars represent standard errors.

either to avoid the garden-path or to recover from it (our experiments do not allow us to distinguish between these two possibilities).

Experiment 4

The final experiment was designed to examine whether the cueing effect found in Experiment 3 could be attributed specifically to the disfluency *uh* or whether any interruption that occurred in the appropriate position might elicit the same effect. Recall that in Experiment 2, environmental noises were capable of producing the head noun position effect. This suggests that if environmental noises and disfluencies both cue the syntactic parser, then the parser is sensitive not to the presence of disfluencies such as "uh" and "um" per se, but rather to any sort of interruption regardless of what it sounds like. To examine this possibility, we assessed the ability of disfluencies (*uh*) and environmental sounds to cue upcoming structure. We once again used the coordination ambiguity and placed disfluencies and noises in the good and bad signal positions from Experiment 3.

If the cueing effect is specific to *uh*, we should see the same pattern of results for the disfluency conditions as we observed in Experiment 3, but no difference between the two environmental noise conditions. On the other hand, if the parser is concerned only with the presence of any sort of interruption to the linguistic input, we might expect to see both conditions show improved parsing (as measured by judgments of grammaticality) in the good signal compared to the bad signal conditions. Such a result would be compatible with the fact that people are somewhat idiosyncratic in their choice of disfluency item, and therefore there exists a wide variety of disfluency types (e.g., "uh"s, "um"s, "er"s, and so on; e.g., Reiger, 2001).

Method

Subjects

Twenty five undergraduate students from Michigan State University participated in this experiment in exchange for research credit in a psychology class. The subjects were all native speakers of English and were naïve to the purposes of the experiment. None had participated in any of the other experiments reported in this paper.

Materials

The stimuli for this experiment were created as follows. The experimental sentences with disfluencies used in Experiment 3 were first duplicated. One of each pair was left unaltered, and these unchanged stimuli made up the disfluency conditions. The other of the pair was changed: The disfluency was removed and replaced with the same environmental noises used in Experiment 2 (cats meowing, dogs barking, telephones ringing, doorbells ringing, sneezing, coughing). The procedure for removing the disfluencies and replacing them with noises was the same as in Experiment 2. Once again, subjects did not find the stimuli unusual apart from the presence of the environmental noises. Twenty coordination ambiguity sentences were included in this experiment (see Appendix). In addition to the 20 experimental sentences, 20 unambiguous grammatical filler sentences and 20 ungrammatical filler sentences were included. The filler sentences were identical to those in Experiment 3, except that half of the disfluencies in the filler sentences were also replaced with environmental noises. The sentences were pseudorandomized for each subject so that no two experimental sentences would occur in succession. Each subject heard only one version of any particular sentence, but heard each of the five conditions an equal number of times. The experiment was run using SuperLab Pro.

Procedure

The procedure for this experiment was identical to that of Experiment 2.

Design and analysis

A 2×2 ANOVA with position (good signal position or bad signal position) as the first variable and intervening material (disfluency or modifiers) as the second variable was conducted. In order to compare the baseline condition to the other conditions, we conducted a one way ANOVA with condition as the variable. In addition to these ANOVAs, t tests were used to compare the two disfluency conditions and the two noise conditions. The proportion of grammatical judgments was the dependent measure in this experiment. Because proportions exceeded 80% in some conditions, the analyses were conducted on arc sin transformed scores (Winer, 1971).

Results and discussion

As in Experiment 3, mild ceiling effects occurred (all conditions were judged grammatical over 80% of the time), and this happened because the coordination ambiguity structure does not produce as strong a garden path as does the subordinate-main structure. The results of the experiment are illustrated in Fig. 6. Main effects of position (by subjects $F_{1,24} = 7.77$, $p < .02$; by items $F_{1,19} = 2.01$, $p > .1$) and material (by subjects $F_{1,24} = 11.01$, $p < .01$; by items $F_{1,19} = 3.36$, $p = .09$) were significant by subjects but not quite by items (Fig. 6). There was no interaction between the variable disfluencies versus noises and the position of the intrusion. The bad signal disfluency condition (83% judged grammatical) was judged grammatical less often than the

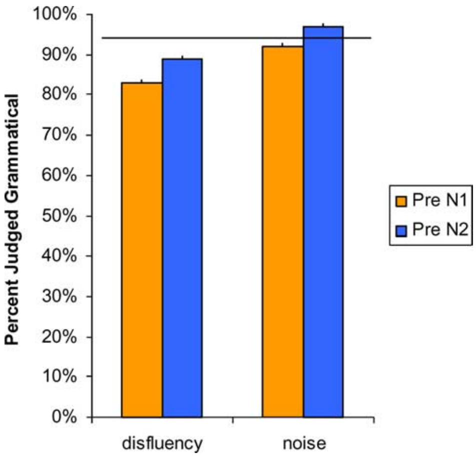

Fig. 6. Percentage of sentences judged grammatical for the coordination ambiguity structure in Experiment 4. The horizontal black line represents the baseline plain ambiguous noun phrase condition. Error bars represent standard errors.

good signal disfluency condition (89% judged grammatical). Likewise, the bad signal noise condition (92% judged grammatical) was judged grammatical less often than the good signal noise condition (97% judged grammatical). The plain ambiguous noun phrase used as a baseline (the horizontal line in Fig. 6) was judged grammatical 94% of the time.

The overall relationship found in the current experiment between the placement of an interruption in an utterance, no matter whether it was a disfluency or an environmental noise, and the subsequent judgment of that utterance's grammaticality indicates that the parser can use the presence of an environmental noise to predict upcoming syntactic structures, just as it uses ''natural'' disfluencies such as *uh*. As Experiment 3 shows, modifiers do not serve this same cueing function. Thus, it seems that the head noun position may be caused by different mechanisms in the cases of modifiers (as modifiers do not seem to signal the parser to expect a major constituent) and disfluent interruptions. The current experiments do not allow us to speculate with any certainty on the mechanism at work when modifiers are present (but see Ferreira & Henderson, 1991, 1998, for discussion). Interruptions, on the other hand, seem to be used as cues to upcoming structure when they occur at appropriate locations.

General discussion

An important principle of parsing that we have exploited in this research on disfluencies is the head noun position effect (Ferreira & Henderson, 1991, 1998). The reanalysis of garden-path sentences is much harder when the ambiguous noun phrase is lengthened with postnominal modifiers. As a first step to determining whether disfluencies might affect processes of syntactic analysis, we wanted to determine whether the same sort of effect might occur when disfluencies occurred in the same position as modifiers. Thus, the first question we asked about disfluencies was simply this: Is it possible to obtain the head noun position effect with disfluencies rather than with modifiers? When a person hears *while the boy scratched the dog* and then one or two ''uh''s, the parser has presumably attached *the dog* as object of *scratched*. The disfluencies come after the head noun of the ambiguous phrase, and prior to the onset of the disambiguating word. The results from the first experiment show that the head noun position effect is produced both when people encounter disfluencies and when they encounter modifying words in the same position. This finding is significant, because it is (to the best of our knowledge) the first demonstration that disfluencies systematically influence the operation of the parser.

We then asked whether we could conclude that disfluencies and words have the same effect—that is, that when disfluencies occur in a location that causes the parser to remain committed to the wrong analysis longer, the parser has more trouble revising that incorrect structure. It might appear obvious that the answer is yes, but we determined that there are in fact two possible explanations of our findings from the first experiment. First, it is possible that the same mechanism is responsible for the head noun position effect in both the disfluency and modifier conditions, and that this mechanism is somehow related to the fact that both disfluencies and modifiers are produced by the speaker's vocal tract. The alternative explanation is that the parser makes use of co-occurrences between disfluencies and structures and uses disfluencies as one cue to the proper attachment of upcoming constituents, and that words cause the head noun position effect by some other mechanism.

This alternative explanation suggests that it is just a coincidence that words and disfluencies produced the same pattern of results in the first experiment. The argument goes as follows: Disfluencies tend to cluster around the beginnings of complex constituents (e.g., clauses; Clark & Wasow, 1998; Ford, 1982; Hawkins, 1971). In the condition where the head noun occurs late in the phrase, the disfluency comes right after the determiner that begins a new clause. Thus, the disfluency predicts a clause and the parser might make use of that information. If it does, subjects will have a cue to the correct parse and so they will be more likely to judge the sentence to be grammatical. When the head occurs early, before the disfluency, the disfluency comes after what seems to be the object of the first clause. The parser might then assume that the disfluency was caused by the speaker's planning of the second clause. That assumption, of course, is incorrect, because the noun phrase that has been attached as object of the first clause is actually the subject of the second. Therefore, the disfluency is misleading, the sentence overall is more difficult to process, and the sentence is less likely to be judged grammatical.

To distinguish between these possibilities, we assessed whether a non-speech interruption could elicit the head noun position effect. To do so, we could not simply leave silence in the location of the disfluencies, because silence could be interpreted as a pause produced by the speaker for the same reason that he or she might say "uh." We decided to try an unusual approach: We replaced the disfluencies with environmental sounds such as telephones ringing, cats meowing, and people sneezing. These sounds are not produced by speakers and are not under their control, and therefore they should not be parsed. The experiment had the same design as the ones described previously, and we found the same results: Sentences were judged grammatical about equally often in the Plain noun phrase and Head-Late conditions, and much less often in the Head-Early condition. Performance was identical with modifying words and environmental sounds. Therefore, we can rule out the hypothesis that the head noun position effect is elicited solely by the presence of speechlike material between the head noun and disambiguating verb. The effect may be due to delay (or some other such mechanism; Ferreira & Henderson, 1991, 1998) in all three cases, but as later experiments showed, this cannot be the whole story.

In the third experiment, we focused on whether disfluencies might serve as either good or bad cues to the syntactic structure of the sentence (assuming that the language comprehension system is sensitive to the tendency of disfluencies to co-occur with the boundaries of major syntactic constituents). We found that disfluencies in positions that made them helpful cues to the ultimately correct structure caused sentences to be judged grammatical more often than even those sentences with no disfluencies (the simple noun phrase condition). Disfluencies in positions that could be considered bad cues, on the other hand, decreased the proportion of sentences that were judged grammatical by listeners. Modifiers showed no such pattern. We took this as evidence that interruptions consisting of disfluencies can be used by the parser as signals of syntactic structure and that modifiers affect parsing by some other mechanism.

The final experiment indicated that an environmental noise in a position that was either a good or bad cue to upcoming structure had much the same effect as a disfluency. This finding, along with those from the other experiments, indicates that disfluent interruptions and modifying adjectives might elicit the head noun position effect for different reasons. Recall that the modifiers in Experiment 3 did not show the same effect of location found in disfluencies, while the environmental noises in Experiment 4 did. This finding raises the possibility that the head noun position effects for disfluencies in Experiment 1 and noises in Experiment 2 may be due in large part to the mechanism suggested in Experiments 3 and 4: namely, the parser making use of the statistical co-occurrence of disfluencies and the initiation of complex constituents (Clark & Wasow, 1998). Because disfluencies, unlike modifiers, do not require the addition of material to the syntactic structure currently being constructed, the parser is faced with a slightly different situation than when it encounters modifying words between the head noun of the ambiguous phrase and disambiguating verb. If the parser were sensitive only to the presence of an interruption, and insensitive to the content of that interruption, it might be able to make predictions about structure when any interruption occurred, regardless of the actual content of that interruption. Such a finding might account for the parser's ability to deal with large variations in disfluency terms. The mechanism responsible for the head noun position effect with modifying words may still be at work to some degree in the disfluent and noise conditions; but our contention is that such a mechanism is superceded by the use of an interruption as a signal to upcoming structure.

These results also help tie together the disparate results discussed concerning the filtering of disfluencies. The finding that listeners make use of disfluent interruptions might lead one to expect that they will recognize disfluencies in a sentence when they occur (Lickley & Bard, 1998). But as the form of interruptions may not be important to the comprehension system, it is not surprising that listeners have difficulty recalling them verbatim (Lickley, 1995; Lickley & Bard, 1996; Martin & Strange, 1968). Moreover, the displacement of disfluencies to clause boundaries may be the result of the cueing effect we have discovered: Because the interruption causes the parser to predict a major constituent such as a clause, we might expect the interruption to be perceived to occur at the boundary of such constituents. Thus, disfluencies may ultimately be "filtered" in the sense that they do not become part of the sentence's final representation; but because the presence of an interruption may affect the syntactic parser by causing it to predict a particular type of constituent, the disfluency may still have influenced language comprehension processes. And any such effect would be in addition to the potential influence of disfluencies on higher level metacognitive judgments such as those described by Smith and Clark (1993) and Brennan and Williams (1995).

Because this is a first step in studying the effect of speech disfluency on syntactic parsing, we must qualify our conclusions to some degree. The disfluencies that occurred in our stimuli were purposely produced to be noticeable. For these first studies, we wanted to ensure that subjects would not miss them. While this is certainly ecologically valid (speakers can and do produce very long and noticeable filled pauses), we do not know whether these same results would be obtained if the disfluencies were much shorter (overall, the duration of the disfluencies was twice as long as the duration of the modifiers). It may also ultimately be relevant that the disfluencies were "staged" rather than being either elicited from naïve speakers via a language production task or selected from a corpus of natural speech. In addition, our stimuli included interruptions in only one location per sentence. Therefore, the signal the disfluency provided was very clear, as was the signal provided by the environmental noises.

Finally, only one kind of filled pause disfluency, the sound *uh*, was examined in this study. We still do not know how other filled pauses might behave. For example, Smith and Clark (1993) have suggested that *um* is a signal of greater difficulty in planning during production compared with "uh," and so, on this view, *um* might be an even better cue to a complex syntactic structure. On the other hand, if the parser is not concerned with details of the interruption, as the results of Experiment 4 suggest, then listeners might not distinguish between "uh" and "um." Another interesting question concerns the way the comprehension system handles repeats and repairs (e.g., *While the man hunted the deer—the rabbit ran through the woods*). Given that these introduce lexical content that must ultimately be removed from the representation, it seems likely that they present the parser with even more interesting challenges.

Appendix

For Experiments 1 and 2, all experimental sentences are included. The * position could be replaced with either the two modifiers following the sentence, or by the disfluency "uh uh." The ~ position could be replaced by "that was" and the second of the two modifiers, or by the disfluency "uh uh." Only one of the * and ~ positions occurred in any one sentence. A baseline condition with no additional material was also used. For Experiment 2, the * position could be replaced with either the two modifiers following the sentence, or by an environmental noise (e.g., a dog barking). The ~ position could be replaced by "that was" and the second of the two modifiers, or by an environmental noise.

For Experiments 3 and 4, all experimental sentences involved the coordination ambiguity. Either the + or the * position could be replaced with either the two modifiers following the sentence, or by the disfluency "uh uh" in Experiment 3. In Experiment 4, either the + or the * position could be replace with either the disfluency "uh uh" or by an environmental noise. Only one of the + and * positions occurred in any one sentence. A baseline condition with no additional material was also used.

Filler sentences are available from either author upon request.

Experimental sentences (coordination structure; Experiments 1–4)

1. Sandra bumped into the + busboy and the * waiter ~ told her to be careful (short and pudgy).
2. Susan punched the + musician and the * conductor ~ went after her (tall and skinny).
3. Wayne defended the + secretary and the * boss ~ thought they were involved (stout and nervous).
4. Phyllis talked to the + mailman and the * neighbors ~ began to gossip (old and lazy).
5. The owner negotiated with the + coach and the * players ~ received raises (rich and well-known).
6. Ken trusted the + builder and his * workers ~ did good work (quick and agile).
7. The actress flirted with the + photographer and the * writer ~ got angry (bald and bearded).
8. Oscar haggled with the + owners and the * real estate agent ~ got upset (glum and cautious).
9. James idolized the + scientist and the * astronaut ~ was very impressed (stern and lonely).
10. Madonna joked with the + players and the * media ~ were all over the place (loud and important).
11. Pamela hit the + fireman and the * policeman ~ had to restrain her (kind and friendly).
12. Dennis kicked the + senator and the * clerk ~ ran to get the Secret Service (mild and pleasant).
13. Bob debated with the + host and the * movie critic ~ called for a truce (young and well-dressed).
14. Audrey scratched the + accountant and the * lawyer ~ advised him to sue (tough and handsome).
15. Jason murdered the + teenager and his * parents ~ became crazy (poor and tired).
16. Samuel kissed the + model and the * actress ~ stormed out in a big huff (blonde and blue-eyed).
17. Harriet missed her + mother and her * brother ~ had to console her (shy and quiet).
18. Axel envied the + athlete and the * singer ~ suspected foul play (wild and flashy).
19. Andrea telephoned the + travel agent and the * assistant ~ called her back (calm and courteous).
20. Kevin fed the + cat and the * parakeet ~ began to chirp loudly (green and orange).

(Experimental sentences (subordinate-main structure; Experiments 1–2, 4)

1. While the man hunted the * deer ~ ran into the woods (brown and furry).
2. While the soldiers fought the * battle ~ ended in defeat (loud and terrible).
3. After the clown tripped the * woman ~ laughed loudly (blonde and blue-eyed).
4. While the girl watched the * television ~ exploded (old and blurry).
5. After the manager finished the * meeting ~ was tolerable (long and boring).
6. After the fire burned the * ranger ~ surveyed the damage (tall and handsome).
7. After the kidnappers returned the * princess ~ cancelled the party (shy and modest).

8. While the ballplayer scratched the * coach ~ yelled directions (bald and pudgy).
9. After the battleship sank the * rowboat ~ radioed for help (red and yellow).
10. After the agents stopped the * car ~ disappeared into the darkness (strange and sinister).
11. After the teacher left the * class ~ broke a window (wild and noisy).
12. While the president prepared the * speech ~ was read without him (short and sincere).
13. After the warriors attacked the * town ~ was evacuated immediately (quaint and sleepy).
14. After the janitor dusted the * clock ~ fell off the shelf (rare and valuable).
15. While the couple hugged the * guests ~ ate the food (young and well-dressed).
16. While the farmer shaved the * sheep ~ trampled the garden (white and fluffy).
17. After the lawyer moved the * box ~ was still in the closet (flat and heavy).
18. After the students protested the convention relocated to another state (large and important).
19. While the general saluted the flag fluttered in the breeze (bright and colorful).
20. While the child studied the suspect sat in the waiting room (dark and scary).
21. Although the cashier understood the woman explained again anyway (mean and angry).
22. While the secretary answered the door opened silently (big and wooden).
23. While the police investigated the family ate their dinner (rich and powerful).
24. While the father lectured the children punched each other (tough and violent).
25. While the customer ordered the burger warmed under the heat lamp (hot and spicy).
26. While the artist painted the king paced nervously (proud and majestic).
27. While the passengers pushed the bus driver told them to sit down (cold and hungry).
28. While the boy bathed the cat caught and mouse (small and nervous).
29. While the chef cooked the chicken pecked at the grain (plump and well-fed).
30. While the priest hid the fugitive emptied the fridge (armed and dangerous).

References

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

Bortfeld, H., Leon, S. D., Bloom, J. E., Schober, M. F., Brennan, S. E. (1999). *Which speakers are most disfluent in conversation and when?* Paper presented at the Disfluency in Spontaneous Speech ICPhS Satellite Meeting, Berkeley, CA, July.

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.

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

Clark, H. H. (1997). Dogmas of understanding. *Discourse Processes, 23*, 567–598.

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

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

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

Dollaghan, C., & Campbell, T. (1992). A procedure for classifying disruptions in spontaneous language samples. *Topics in Language Disorders, 12*, 56–68.

Ferreira, F., & Henderson, J. M. (1991). Recovery from misanalyses of garden-path sentences. *Journal of Memory and Language, 25*, 725–745.

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.

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.

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, 34*, 709–738.

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

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

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.

Levelt, W. J. M. (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). Foris: Dordrecht.

Lickley, R. J. (1995). Missing disfluencies. In *Proceedings of international congress of phonetic sciences* (Vol. 4, pp. 192–195). Stockholm.

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

Lickley, R. J., & Bard, E. G. (1996). On not recognizing disfluencies in dialogue. In *Proceedings of the international conference on spoken language processing 1876–1880*. Philadelphia.

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

Martin, J., & Strange, W. (1968). The perception of hesitation in spontaneous speech. *Perception and Psychophysics, 53*, 1–15.

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

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

Reiger, C. (2001). *Idiosyncratic fillers in the speech of bilinguals*. Paper presented at the Disfluency in Spontaneous Speech ISCA Tutorial and Research Workshop, Edinburgh, Scotland, August.

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.

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

Shriberg, E. E., Bear, J., & Dowding, J. (1992). Automatic detection and correction of repairs in human–computer dialog. In M. Marcus (Ed.), *Proc. DARPA Speech and Natural Language Workshop* (pp. 419–424). Los Altos, CA: Morgan Kaufmann.

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

Traxler, M. J., Bybee, M. D., & Pickering, M. J. (1997). Influence of connectives on language comprehension: Eye-tracking evidence for incremental interpretation. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 50A*, 481–497.

Winer, B. J. (1971). *Statistical principles in experimental design* (2nd ed). New York: McGraw-Hill.