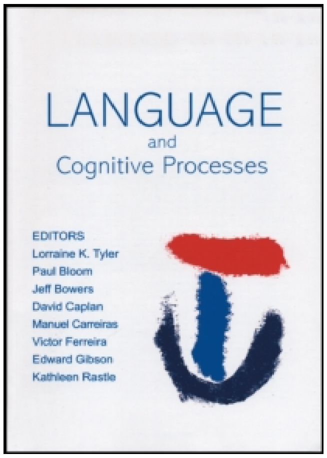

Language and Cognitive Processes
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713683153

Lingering effects of disfluent material on comprehension of garden path sentences

Ellen F. Lau [a]; Fernanda Ferreira [a]
[a] Department of Psychology and Cognitive Science Program, Michigan State University. East Lansing, MI. USA

To cite this Article: Ellen F. Lau and Fernanda Ferreira , 'Lingering effects of disfluent material on comprehension of garden path sentences', Language and Cognitive Processes, 20:5, 633 - 666
To link to this article: DOI: 10.1080/01690960444000142
URL: http://dx.doi.org/10.1080/01690960444000142

LANGUAGE AND COGNITIVE PROCESSES
2005, 20 (5), 633–666

Psychology Press
Taylor & Francis Group

Lingering effects of disfluent material on comprehension of garden path sentences

Ellen F. Lau and Fernanda Ferreira
Department of Psychology and Cognitive Science Program, Michigan State University, East Lansing, MI, USA

In two experiments, we tested for lingering effects of *verb replacement* disfluencies on the processing of garden path sentences that exhibit the main verb/reduced relative (MV/RR) ambiguity. Participants heard sentences with revisions like *The little girl chosen, uh, selected for the role celebrated with her parents and friends*. We found that the syntactic ambiguity associated with the reparandum verb involved in the disfluency (here *chosen*) had an influence on later parsing: Garden path sentences that included such revisions were more likely to be judged grammatical if the reparandum verb was structurally unambiguous. Conversely, ambiguous non-garden path sentences were more likely to be judged *un*grammatical if the structurally unambiguous disfluency verb was inconsistent with the final reading. Results support a model of disfluency processing in which the syntactic frame associated with the replacement verb ‘‘overlays’’ the previous verb's structure rather than actively deleting the already-built tree.

INTRODUCTION

Because of the relative complexity of working with auditory materials and data, the bulk of psycholinguistic research on human language comprehension has been done with written materials. When we use the results of

Correspondence should be addressed to Fernanda Ferreira, Cognitive Science Program and Department of Psychology, Michigan State University, East Lansing, MI 48824-1117, USA. Email: Fernanda@eyelab.msu.edu.

This research was supported in part by grants from the National Science Foundation (BCS-9976584, DGE-0114378) and a Michigan State Department of Social Science Dean's Assistantship awarded to Ellen Lau. The authors thank Karl Bailey, Cristina Schmitt, and Jim Zacks for comments on the manuscript.

The first author is now at the Department of Linguistics, University of Maryland, College Park, MD, USA.

http://www.tandf.co.uk/journals/pp/01690965.html DOI: 10.1080/01690960444000142

such research to develop theories about the nature of the human language processor, we usually make the implicit assumption that language processing works similarly in all modalities. Although this assumption is valid and useful for the most part, we too often forget that speech possesses a number of characteristics that are not found in written language. The prosodic contour of spoken sentences is perhaps the most widely discussed example of late, but speech differs from text in other significant ways as well: These include the speaker's control of the presentation rate, the impossibility of listener-backtracking, and a relatively high occurrence of the errors and online corrections that fall under the heading of *disfluency*. After many years of landmark work on written-language processing, psycholinguists have increasingly begun paying more attention to such spoken-language phenomena. This turn of events is desirable in part because learning more about speech comprehension is important and interesting in its own right—after all speech, not writing, is the original mode of human communication—but also because the findings of such research will not necessarily be limited in scope to spoken language. One example of the potentially broad-ranging impact of such research has been seen in recent work arguing that influences of prosodic structure on the interpretation of ambiguous sentences may make themselves felt even in written material, an effect now known as ''silent'' prosody (Bader, 1998; Fodor, 1998).

Although a considerable amount of work has focused on the effects of prosodic representations on sentence comprehension, the corresponding literature on disfluency is relatively sparse. In the set of experiments described here, we examine the effects on sentence processing of one type of disfluency, *replacement,* as a way of addressing issues concerning the representation and storage of disfluencies in general as well as the sentences that contain them. The work also has important implications for basic issues in sentence comprehension. Much of the work in the substantial literature on garden path sentences has been focused on how the parser recovers from an initially incorrect syntactic/semantic interpretation and ''reanalyzes'' the sentence (e.g., Christianson, Hollingworth, Halliwell, & Ferreira 2001; Frazier & Rayner, 1982). Significantly, a listener encountering a repair-type disfluency has much the same task: He or she must recover from an interpretation built on the basis of information that turns out to be unintended and incorrect with respect to the sentence's ultimate form.

In the two experiments we describe in this paper, we examine sentences with verb replacement disfluencies like *The little girl chosen, uh, selected for the role celebrated with her parents and friends*. If at the replacement verb the parser strikes out all of the structure associated with the disfluency word *chosen* and processes *selected* with a 'clean slate', we would expect the material following the disfluency and repair to be processed as if no disfluency had occurred. If at the replacement verb the disfluency word *chosen* and the structure associated is actually *inhibited* to prevent conflict with the new material, we might expect the dispreferred RR-reading of *selected* to become even more difficult. However, if the previously built sentence structure and/or structure associated with the disfluency item itself remain activated, we predict that the processing of the subsequent material will be affected; in particular, that the disfluency will make the garden path sentence above easier to process.

Previous work examining disfluency in sentence processing

Disfluency in natural speech is widespread. Most estimates put the number at somewhere around 6 disfluencies per 100 words, where disfluencies include repetition, self-interruption, self-correction, and expressions such as *um* and *I mean* (Fox Tree, 1995). From both the speaker and listener standpoint, disfluencies tend to be viewed as deviations from an ''ideal'' speech act (Clark, 1996). In this sense, we often characterise disfluency roughly as an indication of mistakes/failures in the production system (although it may simultaneously serve certain positive discourse functions, e.g., holding the floor while recovering from production failure).

Assuming this view of disfluencies as ''mistakes'', one hypothesis might be that disfluencies are simply filtered out by the parser and that they play no interesting role in comprehension. Although not specifically addressing the effects of disfluency in sentence processing, the findings of Lickley and colleagues are consistent with this idea (Lickley, 1995; Lickley & Bard, 1996), showing that listeners generally cannot correctly identify the location of filled pause disfluencies (e.g., *uh, um*) in sentences they have just heard, and often do not report them at all. The vast majority of natural language processing systems developed to deal with disfluencies also assume a type of filtering architecture (e.g., Charniak & Johnson, 2001; Stolcke & Shriberg 1996; although cf. Core & Schubert, 1999). But while a machine parser armed with a simple disfluency-filtering mechanism may function adequately,[1] several recent studies on disfluency demonstrate convincingly that the human parser does not work this way.

[1] Note that the results of Bailey and Ferreira (2003) imply that such a parser's performance would nevertheless be improved further if it did make use of disfluency information.

Filled pauses in comprehension

First, although disfluencies such as filled pauses which carry little semantic content would seem to be the least disruptive kind and thus the least likely to affect comprehension, complex effects of these items have been demonstrated on various levels of processing. In a colour-matching computer task, Brennan and Schober (2001) found speedier response times to sentences with replacement disfluencies when *uh* or *um* was also part of the repair segment, and they hypothesised that filled pauses may act as helpful cues that repair is needed. Brennan and Williams (1995) found that listeners could correctly determine the degree of confidence speakers felt in their responses to queries merely by tracking the *type* of filled or unfilled pause present (i.e., *uh* versus *um* versus unfilled pauses); these results suggest that disfluencies also signal pragmatic information about the message itself.

Most recently, work by Bailey and Ferreira (2003, in press) suggests that filled pause disfluencies also influence syntactic analysis in several ways. For example, just by filling up extra time in a given portion of the sentence, filled pauses increase the amount of time that the parser is committed to a current interpretation. In previous work, Ferreira and Henderson (1991a) showed the existence of what they termed a Head Position Effect (HPE) in parsing, finding that recovery from a garden path is more difficult when extra lexical material (such as modifiers) intervenes between the head of a temporarily ambiguous phrase and the disambiguation point (*While the man hunted the deer ran through the woods* vs. *While the man hunted the brown and furry deer ran through the woods*). Ferreira and Henderson proposed that the extra material lengthens the amount of time that the incorrect reading is maintained, thereby increasing the strength of the commitment and making a reversal of this commitment more difficult (see Ferreira and Henderson, 1991b for more detailed discussion of the mechanism). Bailey and Ferreira (2003, in press) predicted that filled pauses would have the same kind of effect as lexical materials. The results of their experiments fulfilled this prediction, showing that sentences with filled pauses in postnominal positions were more likely to be judged ungrammatical than sentences with filled pauses in prenominal positions (*While the man hunted the deer uh uh ran through the woods* vs. *While the man hunted the uh uh deer ran through the woods*).

In further work, Bailey and Ferreira (2003) demonstrated another, related means by which disfluencies can affect syntactic analysis. Because filled pauses are statistically more likely to occur at certain syntactic positions in the parse (namely, before major constituent boundaries), they can act as cues for the listener of upcoming structure. Their results bore out this hypothesis: listeners were more likely to assign phrase boundaries correctly in an ambiguous structure (as measured by grammaticality judgements) when the position of a filled pause was consistent with this assignment, and less likely to make a correct decision when the position of the filled pause was inconsistent with the assignment (the effect was not seen in a control condition with modifiers).

Repetitions and revisions in comprehension

The abovementioned studies suggest that, contrary to a filtering hypothesis, disfluencies such as filled pauses have at least some effect on the linguistic representations built by the listener, even if we do not completely understand the mechanisms by which this occurs. Given the work summarised above on disfluencies with minimal semantic/syntactic content, we might expect more contentful types of disfluency to similarly influence sentence comprehension—perhaps in even more complex ways, given the large amount of structural information that is assumed to be associated with lexical items in major psycholinguistic theories (e.g., MacDonald, Pearlmutter, & Seidenberg, 1994).

The most common type of such ''lexical'' disfluencies—in some sense the baseline case—is *repetition*, in which the speaker interrupts him or herself, returns to some earlier point in the utterance, and repeats the same material (*I'm going to, to Graceland*). Somewhat less common and (theoretically) more difficult for the listener are what we will refer to here as *revisions* (also sometimes called *false starts*, *substitutions, replacements,* or *reformulations*), in which the speaker interrupts him- or herself and returns to an earlier point, but makes a change to the output (*I'm headed for, uh, I'm going to Graceland*). As in this example, filled pauses often precede the repetition or revision. While revisions can involve some repetition of material in addition to the change (*I'm* ... in the previous example), they can also be limited to the changed portion (*I'm going to Mecca, uh, Graceland*). For the purposes of this work, we will refer to this latter type of revision as a *replacement,* because the incorrect material is simply replaced by the correct material.

In order to discuss such disfluencies in more detail, we will at this point introduce some terminology. A representation of how these terms map on to the phenomena in question is given in Figure 1. The *reparandum* refers to the lexical material that ends up being repaired; therefore, this segment will contain at least some of the fluent speech preceding the self-interruption that begins the speaker's correction process. The *edit interval* (also sometimes called the *disfluency interval*) refers to the section between the interruption point and the repair, and often includes pauses or expressions like *uh* or *I mean.* Finally, the *repair interval* refers to the lexical material following the edit interval that is meant to correct or

Repetition	I'm going to,	uh, well,	I'm going to	Graceland
	{-reparandum-}	**{-edit interval-}**	**{-repair interval-}**	
Revision	I'm headed for,	uh, I mean,	I'm going to	Graceland.
	{-reparandum-}	**{-edit interval-}**	**{-repair interval-}**	

Figure 1. Repair terminology.

replace the reparandum; this includes repeated material as well as any new or corrected information (Brennan & Schober, 2001; Nakatani & Hirschberg, 1994).

The key difference between repetitions and revisions, then, is that when listeners encounter a revision, they must get rid of some of the original information present in the reparandum in order to form a correct representation of the utterance. Processing a repetition, on the other hand, requires only a sort of superimposing of identical information. Furthermore, the process of locating the beginning of the reparandum may be more difficult in revisions than in repetitions, where the reparandum can be located simply by identity matching. Although prosodic cues probably play an important role in the processing of disfluency (for a review of such effects, see Nakatani and Hirschberg, 1994), Lickley and colleagues showed that the earliest point at which listeners can use prosody to recognise the onset of disfluency is the edit interval (Lickley, Shillcock, & Bard, 1991; Lickley & Bard, 1992). These results imply that prosody does not reliably mark the reparandum segment; therefore, if the repair interval takes a very different form from that of the reparandum, it may be difficult for the listener to determine where the reparandum begins. Similarly, if there are no prosodic cues, filled pauses, or discourse markers in the edit interval, it may be harder for the parser to recognise that a revision has occurred at all: Compare *I just saw Elvis, Graceland* and *I just saw Elvis, I mean, Graceland*. The former could be part of a coordination structure, an analysis that is disconfirmed only when the sentence ends. In contrast, *fluent* repeated material is rare enough that any repetition of a lexical string can in itself be a cue that a disfluency has occurred, as in *I just saw Graceland, Graceland* (but see Howell and Young, 1991 for evidence that prosodic cues are—perhaps compensatingly—more marked for revisions than repetitions).

For these reasons, we might expect that revisions would result in more processing difficulty than repetitions, and in fact this is exactly what Fox Tree (1995) found. Fox Tree used a word-monitoring task with fluent and disfluent sentences and showed that while response latencies to revisions were longer than fluent controls, response latencies to repetitions did not differ. She argued that this asymmetry resulted from difficulty in locating the repair site of revisions rather than from the need to delete old material, and presented several post-hoc analyses consistent with this claim. These analyses included the lack of a correlation between response time and the syntactic/semantic discrepancy between the reparandum and repair, as well as reduced difficulty for revisions that occurred at the beginning of sentences where location of repair would likely be simple. However, while her data are consistent with a story in which localisation of the repair site is a main source of difficulty in processing revisions, Fox Tree's experiments did not show explicitly that the need to delete information was not a factor.

Brennan and Schober's (2001) results, on the other hand, do suggest that the cost of deletion/inhibition plays a part in the increased difficulty of revisions. Brennan and Schober presented disfluent commands to participants as part of an object-selection task and found that participants had an easier time recovering from error-truncated disfluencies (*Move to the yel- purple square*) than non-truncated disfluencies (*Move to the yellow-purple square*). They also found that in both mid-word and between-word disfluency conditions, error rates in the response were positively correlated with the length of the reparandum heard, and reaction times were negatively correlated with the length of the edit interval. In other words, if more wrong information was heard, recovery was more difficult; if the (filled) pause in-between the false information and the correct information was longer, recovery was facilitated. These results support a model in which the time benefit in processing revisions is a result of limitations on processing resources that must (1) use cues to realise a disfluency has occurred, and (2) inhibit some or all of the original representation: If listeners have a longer edit interval in which they can accomplish these things without having to process new contentful information, recovery will be more successful than in the absence of a pause when they must divide their resources between the tasks. Further, if the interruption of the word limits its activation, or cues active inhibition, its potential interfering effect on the new word should be diminished. Of course, an alternative explanation that does not reference processing limitations would be that a longer pause is simply a better cue to the listener that a disfluency has occurred and thereby speeds up the recognition process, leading to facilitated comprehension. Thus, future empirical work is needed to determine more conclusively the source of the relative difficulty of revisions.

Potentially lingering effects of structural information in revisions

Although the ''costly-inhibition'' interpretation of Brennan and Schober's (2001) results remains a matter of some debate, their study inspired us to ask more detailed questions about the mechanisms the parser uses to deal with the lexical material involved in disfluencies such as revisions. How complete is the deletion involved in the revision: Is the old material completely deleted or inhibited *à la* the filtering hypothesis, or is some of the semantic/syntactic information preserved—either by design or as a byproduct of limited processing resources? Also, are structure-building and interpretative decisions made on the basis of the replaced material also reversed when the parser recognises the occurrence of a revision? We suspected that the answer to this latter question would be ‘‘no’’, and that both lexical material from the reparandum and structural material from the initial parse would still be active enough to have effects on the revised parse.

Here we set out to answer such questions about the extent of disfluency ‘‘reanalysis’’ by focusing specifically on the processing of verb replacement disfluencies. The reason for this focus is simple: We know that verbs carry a great deal of structural information as part of their lexical entries. In laying out the basic architecture of the constraint-based lexicalist processing model, MacDonald, Pearlmutter, and Seidenberg (1994) popularised the view that the lexical representation of a word includes not only phonological, orthographic, morphological, and semantic information, but also structural information like grammatical features, X-bar structure, and argument structure. Although the predictions made by such models can in theory be applied to the lexical representations of any word category, in psycholinguistics they have been most often examined for the lexical entries of verbs, whose selection and subcategorisation preferences tend to have the most visible effects on processing. Much of this work has confirmed a significant influence of lexical information on the parsing of sentences (e.g., Altmann & Kamide, 1999; McElree & Griffith, 1995; Shapiro, Zurif, & Grimshaw, 1989; Trueswell & Kim, 1998). In addition, tree-adjoining grammars (TAGs) have recently provided a means of making lexical frameworks fairly explicit (Ferreira, 2000; Kim, Srinivas, & Trueswell, 1999).

In the following set of experiments, we look at the effects of this kind of structural information on processing when a verb is part of a replacement disfluency. If thematic roles have already been assigned and syntactic structure has already been built based on the subcategorisation frames associated with this verb, will the roles be unassigned and the structure deleted when the replacement is encountered? We predicted that, on the contrary, some of these parsing consequences of the original material would linger and affect the processing of the edited sentence, in a way analogous to the lingering-interpretation effects shown in garden path sentences by Christianson et al. (2001). In fact, we can almost think of revisions as a kind of garden path, in which reanalysis might or might not be necessary depending on the type of repair that is made by the speaker.

To ensure a mistake-free parse, when listeners realise that a revision has occurred they should not only ‘‘strike out’’ the lexical items that are being replaced by the lexical items in the repair interval; they should also ‘‘strike out’’ structure built and thematic roles assigned on the basis of the original lexical items, because they may not be compatible with the items that replace them. However, all of this ‘‘destruction’’ and ‘‘reconstruction’’ of parts of the analysis will require processing mechanisms and resources to implement them. If the amount of time to recover from the repair is limited—as it usually is during normal conversation—we expect that deletion of the earlier parse will not be complete. Also, it might be that lexically-specific features (e.g., subcategorisation frames/elementary trees) of the materials present in the reparandum will exert a priming effect on the items that replace them, such that they influence which structural/semantic forms of the replacement material are selected by the parser. For both of these reasons, we predicted that we would see effects of information associated with the earlier verb on the parsing of the final, edited sentence.

In our first experiment, we looked at effects of verb replacements on grammaticality judgements for garden path sentences involving the main verb/reduced relative ambiguity. If the syntactic structure building/thematic role assignment set in motion by the replacement verb is not completely suppressed, we expect that sentences with a reparandum verb that unambiguously predicts the correct structure will be more likely to be judged grammatical than those with a reparandum verb that is ambiguous and favours the incorrect structure. The second experiment had a similar design, but crucially in this case we used materials that were disambiguated to the preferred, *non*-garden path reading. This manipulation was designed to see if the disfluency effect was strong enough to elicit a *reverse* garden path effect; by this we mean that the disfluency could cause readers to actually be biased against the main-verb reading that countless studies have shown to be preferred (Ferreira & Clifton, 1986; Frazier, 1978; MacDonald, 1994).

EXPERIMENT 1

For simplicity of exposition, we will define some basic assumptions about the human sentence processor before describing the logic of our design. In the discussion that follows we will assume a serial parser, so that at any point in the sentence the parser is committed to a given analysis, and also that parsing to some extent is predictively incremental, in that the parser can make commitments about structure before the constituents are actually encountered. The model we assume is also constraint-based in the sense that various sources of probabilistic information are used to resolve ambiguities, largely at the lexical level (MacDonald, Pearlmutter, & Seidenberg, 1994; Trueswell, 1996; Trueswell, Tanenhaus, & Garnsey 1994). We make these assumptions for convenience only, and it is important to note that our predictions would be consistent with any parsing framework that relies heavily on lexicalised verb information, thus including most parallel parsing models.

First we will run through an example of how our assumed model handles typical, fluent ambiguous sentences, using a sentence with the main-verb/reduced-relative (MV/RR) ambiguity that we will focus on in these experiments: *The little girl selected for the role celebrated with her parents and friends.* When the parser encounters the verb *selected*, the sentence is ambiguous between a transitive main verb continuation *The little girl selected one piece of candy* and the ultimately correct reduced relative continuation (called a reduced relative because of the absence of a complementizer introducing the relative clause; *The little girl (that was) selected* ...). The relative activation of the two possibilities will be mediated by various types of frequency information. The lexical entry for *selected* is associated with elementary trees corresponding to both the MV and the RR interpretation. In the case of *The little girl selected for the role* ... , a greater frequency is associated with the MV tree. Because the MV interpretation corresponds with thematic assignment of the <Agent> role to the subject, this reading is also encouraged by contextual information: *the little girl* is animate, and agents tend to be animate. These constraints give weight to the initial, incorrect selection of the MV tree into which the initial NP *the little girl* is inserted or, in LTAG terms, *substituted*. When no appropriate object is encountered to substitute into the other argument position of the MV tree, the parser must reanalyse the sentence, resulting in the experience of a garden path (MacDonald, 1994).[2]

Because structural information on the verb plays such an important part in the processing of this kind of sentence, the MV/RR garden path provides a good opportunity to examine our predictions about how the processor deals with the structural information associated with the lexical

[2] The reader may have noticed that the MV/RR garden path sentence presented here does not give rise to the same level of conscious difficulty as the famous example of the same type, *The horse raced past the barn fell*. It has been suggested that the relative ease of garden paths like *The little girl selected for the role celebrated with her parents* results at least partly from the inability of the PP for the role to appear after the verb in a MV structure, leading to earlier disambiguation, as in **The little girl selected for the role*. vs. *The horse raced past the barn* (Gibson 1991; Pritchett 1988). But while the garden path is milder in these cases, such sentences have still been found to be measurably more difficult than their unambiguous counterparts; this will be demonstrated again in the following experiments as a baseline measure.

items that are involved in replacement disfluencies. As described above, the processor proceeds incrementally through the garden path sentence, encounters *selected*, and on the basis of various constraints, selects the MV tree and substitutes the initial NP. However, what if *selected* were part of a replacement, as in *The little girl selected-uh, picked for the role celebrated with her parents and friends*? We can imagine several things the parser might do. The parser might completely delete the old material, discard the entire original parse, and reconstruct the sentence with the new verb *picked* in place of *selected*, by choosing the MV tree for *picked* and substituting *the little girl* into *picked*'s elementary tree. Alternatively, the parser might preserve the structure already built through the combination of elementary trees, and merely replace the phonological and possibly the semantic[3] information from the old verb *selected* with the new verb *picked*.

These alternative mechanisms can be distinguished by comparing this case with one in which the verb in the reparandum (henceforth the *reparandum verb*) has a different set of associated elementary trees than the verb that replaces it (henceforth the *replacement verb*). If a listener hears a replacement in which the reparandum verb is like *chosen* in having only an RR elementary tree (*The little girl chosen, uh selected for the role celebrated with her parents and friends*), the RR structure should start to be filled in as soon as *chosen* is encountered and the initial NP *the little girl* will be substituted into this tree. When the parser encounters *uh, selected*, it must replace *chosen* with *selected* and should delete the structure built and start over, because the information associated with the replacement verb (*selected)* does not allow a choice between the RR structure and the MV structure (and actually favours the latter). In this case processing would proceed normally, and the usual garden path of assuming MV rather than RR structure should occur. If the process of deleting and starting over involves active inhibition of the reparandum material, and if this inhibition is applied to the structural possibilities associated with the lexical items in the reparandum, the RR structure would be even less active than usual, and the garden path should actually be worse. However, if the original structure is not actively deleted from memory and the phonological/ semantic representations associated with *selected* simply replace those associated with *chosen* in the RR structure, we might expect the garden path to be less severe or even not to occur at all.

---

[3] By semantic here, we mean the ‘‘dictionary’’ meaning associated with the wordform and not the associated argument structure requirements. In the cases we will focus on, both the argument structure and the basic meaning of the substituting verbs are fairly similar; thus we make no claims about the possibility that this kind of semantic information from the disfluency lingers as well, although it certainly seems likely.

The former version of the sentence, in which the two verbs have the same structural options (*The little girl picked-uh selected for the role* ...), could conceivably be even *more* difficult than the baseline, fluent garden path case, because the presence of the replacement verb means that the incorrect MV interpretation will persist for a longer period of time, much like the head position effect discussed earlier. In any case, the reparandum verb *picked* should not make resolution of the ambiguity any easier. Therefore, if the structural information associated with the reparandum verb affects parsing, we predict processing differences between these two cases.

In these experiments, we used a simple, end-of-the-sentence grammaticality/acceptability judgement as the measure of garden path difficulty. This measure has been used to obtain reliable differences in many garden path studies, the assumption being that on a certain percentage of trials the parser will fail to recover from the garden path and assign an 'ungrammatical' tag to the sentence; this proportion should then vary with reanalysis difficulty (see discussion in Ferreira and Henderson, 1991a). In the current experiments, participants were asked to make judgements about disfluent sentences as well as fluent ones, raising the possibility that the judgements will be influenced by the fluency of the message as well as the difficulty of the structural analysis. We do report a significant effect of disfluency on the judgements below (although the difference is numerically small), and this suggests that we cannot make direct comparisons between the fluent and disfluent conditions. At the same time, we should be able to safely take any differences between the disfluent conditions themselves to represent effects of ease of processing over and above the mere presence of disfluency.

Method

Participants. Forty-two undergraduates from Michigan State University participated in the experiment in exchange for course credit. All were native speakers of English and had normal hearing.

Materials. The stimuli consisted of 24 experimental items made up of 12 verb triples (Appendix A) and 96 fillers. Each experimental item appeared in one of four variations of a sentence in which the reduced-relative reading was always the correct one. Because the number of common unambiguous past-participle forms in English is limited, we used each of the 12 verb triples in two items. Some of the materials were sentences from MacDonald (1994, Expt 1A), slightly modified for auditory considerations. Each sentence began with a 3-word NP (all including an adjectival modifier), followed by the verb(s), a prepositional phrase, and a short predicate (see 1a–d below).

In the garden path conditions (1b–d), the sentence was completely disambiguated at the second verb,[4] but the sentences were designed in such a way that they were largely disambiguated by the end of the preceding PP.[5]

(1)

a. The little girl chosen for the role celebrated with her parents and friends.
   (fluent, unambiguous)
b. The little girl selected for the role celebrated with her parents and friends.
   (fluent, ambiguous)
c. The little girl chosen-uh selected for the role celebrated with her parents and friends.
   (disfluent, unambiguous reparandum)
d. The little girl picked-uh selected for the role celebrated with her parents and friends.
   (disfluent, ambiguous reparandum)

In the fluent, unambiguous condition (1a), the verb is of an unambiguously past-participle form, making the RR reading the only one possible. In the fluent, ambiguous condition (1b), the verb is ambiguous between a simple past and a past-participle form, resulting in a garden path to the incorrect MV reading. In the disfluent, unambiguous reparandum condition (1c), a replacement disfluency occurs in which the ambiguous replacement verb is preceded by a

[4] In both experiments we inadvertently included a verb (raced) which can appear not only in the transitive MV and RR constructions discussed here, but can also appear intransitively. This may have resulted in a later disambiguation point than other items. Although it is unclear that our main predictions in this case would differ (disambiguation would presumably still happen at the verb), all reported effects remained significant when this item was excluded from analyses.

[5] In general, the PPs following the first verb would preclude an MV reading (**The little girl selected for the role Maria* vs. *The little girl selected Maria for the role*). However, in English, it is possible for a PP to precede a direct object if the direct object is fairly long or ‘‘heavy’’ through a process known as ‘‘heavy NP-shift’’ (*The little girl selected for the role her best friend Maria who had lived next door to her since kindergarten*). Because such constructions are at least somewhat marked, we still believe that disambiguation generally happened at the PP in our material (see MacDonald, 1994). At the same time, we should note that our predictions do not at all depend on whether disambiguation takes place at the PP or at the second verb; in either case, the disfluency takes place before the disambiguating material.

reparandum verb that has an unambiguous past-participle form, like *chosen*. In the disfluent, ambiguous reparandum condition (1d), a replacement disfluency occurs in which the ambiguous replacement verb is preceded by a reparandum verb that is similarly ambiguous between the two structures. The two fluent conditions were included to give a baseline measure of garden path strength over all the verbs. The 96 fillers were balanced for grammaticality, presence of disfluency, and placement of disfluency (on the verb vs. elsewhere). All disfluencies were of the replacement type.

Although each verb triple appeared twice in the experimental items, the materials were manipulated in such a way that there was little within-experiment bias of the structural environment associated with the verb. This was possible because, as in the example above, the verb triples consisted of one unambiguous verb (*chosen*) and two semantically similar ambiguous verbs (*picked* and *selected*). Therefore, as (1a–d) above formed one experimental item, another item was constructed in which *picked* and *selected* switched places in the paradigm. This allowed us to arrange the stimulus lists so that no participant heard a given ambiguous verb in a reduced-relative structure more than once. Not only did this design allow us to double the number of experimental materials, but it provided a chance to balance any inequity in the degree to which the verbs were transitively biased.

The stimuli were recorded by a male native speaker of English, who read the sentences with normal, meaningful prosody. Although some types of garden path sentences can be disambiguated prosodically, we did not expect the auditory nature of our materials to minimise their difficulty. Ferreira, Henderson, Anes, Weeks, and McFarlane (1996) showed MV/RR garden path effects identical to those in written studies using an auditory moving window comprehension task, suggesting that prosody does not play a role in disambiguating this kind of structure. Bailey & Ferreira (2003, in press) describe in more detail the stimulus properties of spoken garden path sentences.

In contrast to Brennan and Schober's (2001) study which used disfluencies elicited from participants doing a task, the disfluencies used in these experiments were artificially created. Although our speaker read the disfluent materials with what seemed to be appropriate prosody, the use of naturally occurring disfluencies would obviously be preferable. However, in contrast to Brennan and Schober's materials, which were simple instructions like *pick up the yellow circle*, our materials were not of a type such that eliciting them naturally from naïve participants would have been feasible. We can note that participants did not report anything strange about the sentences, although they were not specifically questioned about the naturalness of the disfluencies.

To make sure that the kind of disfluency phenomena we examined do in fact occur in real speech, we analysed a set of 1121 disfluency-tagged Switchboard dialogues (comprising 2,247,330 words). We found that replacements made up a small but meaningful percentage of disfluencies (4.2%, or 1970 occurrences out of a total of 47,155 disfluency-tagged phenomena) although, in line with previous counts (Shriberg, 1996), replacements did not occur nearly as often as repetitions (50%, or 23,553 occurrences). Verb replacements, in which one verb was switched with another verb, did happen but were relatively rare (0.7%, or 325 occurrences); these included cases where two different forms of the same verb were switched. Thus in using verb replacements as the focus of our manipulation, we do not assume that listeners must deal with revisions involving these particular categories relatively frequently. We use replacements of verbs rather than other categories simply because verbs have the necessary associated syntactic information to provide a good means of finding out more about disfluency processes in general. What is important for our purposes in this study is that replacements as a broad category are common enough that listeners must have some way of dealing with them.

Procedure. The procedure was the same as that used by Bailey and Ferreira (2003). Before the start of the experiment, participants were given examples of grammatical and ungrammatical sentences. They were also instructed that, as in real life, they should not treat the occurrence of a corrected error as an example of ungrammaticality, but should make their judgements based on the entire sentence. Eight practice trials were presented before the experiment to allow participants to become familiar with the task. At the beginning of each trial, participants pressed a mouse button to hear the sentence. Once the sentence was complete, they pressed either the right or left mouse button, for grammatical or ungrammatical, respectively. Judgements were recorded by SuperLabPro experimental software.

Results

The data from Experiment 1 are presented in Figure 2. Data from two participants were not included because their accuracy on grammatical fillers fell below our 90% accuracy cutoff. Because proportions were overall close to 1, all analyses were done on arcsine transformed scores. A 2×2 within-subjects ANOVA was done using as factors the presence/absence of disfluency and first-encountered verb type. The (fluent) unambiguous and the unambiguous disfluency conditions were combined under the verb type factor because the first verb encountered was

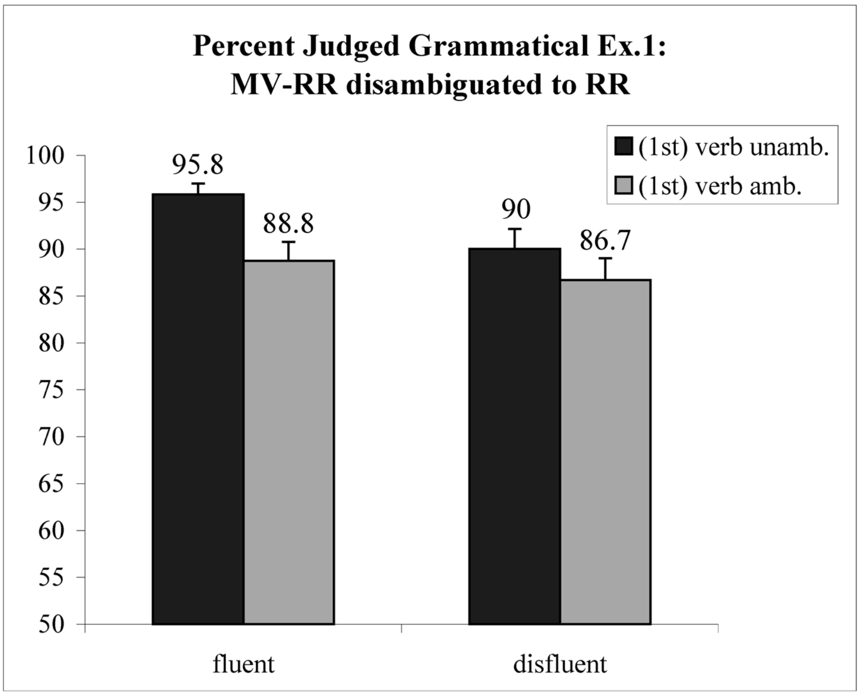

Figure 2. Percentage judged grammatical for conditions described in (1). Error bars represent standard errors.

unambiguous, although the disfluent condition became ambiguous again when the verb was replaced. Under our hypothesis, encountering a disfluency with an unambiguous verb in a globally ambiguous sentence should be similar to encountering an unambiguous verb in a fluent sentence; thus we predicted a main effect of verb type and no interaction. Planned comparisons were also made within the fluent and disfluent conditions respectively: in the fluent condition as a test for the baseline garden path, and in the disfluent condition as a test for our hypothesised influence of disfluency verb type on sentence processing.

Grammaticality judgements were uniformly high (90% averaging over the experimental items), which is expected given that the MV/RR garden path in these materials is a relatively mild one (see footnote 2). Grammaticality judgements for grammatical and ungrammatical fillers were at predictable extremes: 98% and 23%, respectively. It is immediately clear that the strongest version of our prediction—that a potentially helpful disfluency in a garden path sentence might make an ambiguous sentence as easy as an unambiguous sentence—was not met. However, a 2×2 ANOVA with verb ambiguity and disfluency as factors revealed a significant main effect of verb type; by subjects, $F(1, 39) = 8.529$, $p < .01$; by items, $F(1, 23) = 8.285$, $p < .01$, and no significant interaction, by subjects, $F(1, 39) = 1.735$, $p = .196$; by items, $F(1, 23) = .051$, $p = .824$. The ANOVA also showed a main effect of disfluency that was marginally significant by subjects, although not by items; by subjects, $F(1, 39) = 3.836$, $p = .057$; by items, $F(1, 23) = 2.375$, $p = .137$; we note, however, that the items analyses were run with a much lower n than the subject analyses.

The significant main effect of verb type and the lack of an interaction suggest that the fluent and disfluent sentences patterned together with respect to the crucial verb, although the disfluent sentences may have been somewhat more likely to be judged ungrammatical than the others. This suggestion was confirmed by the results of planned comparisons within the fluent and disfluent conditions. A paired-sample *t*-test on the two fluent conditions (garden path vs. non-garden path) was significant; by subjects, $t(39) = 5.01$, $p < .01$; by items, $t(23) = 2.461$, $p < .05$, demonstrating the existence of the baseline garden path effect in our auditory materials. Crucially, this effect was replicated in a paired-sample *t*-test on the two disfluent garden path conditions (potentially helpful disfluency and neutral disfluency), which was significant by subjects and marginally significant by items; by subjects, $t(39) = 2.475$, $p < .05$; by items, $t(23) = 2.009$, $p = .056$. Thus, even though the final, corrected versions of the sentences in these two disfluency conditions were identical, judgements were significantly higher for sentences in which information in the "mistake" actually predicted the non-preferred RR continuation.

DISCUSSION

The results of Experiment 1 show that the structural possibilities associated with the reparandum verb in a replacement disfluency exert a small, but measurable, positive effect on the parsing of the remainder of the sentence. If the reparandum verb's form unambiguously indicated a RR-structure for the sentence, it was more likely that participants would form a grammatical interpretation for the corrected-to-ambiguous sentence when compared with cases in which the reparandum verb was also ambiguous. These results contrast with the predictions of both the complete-deletion view (prediction: no difference in disfluency conditions) and the active-inhibition view (prediction: *increased* difficulty in the unambiguous disfluency case). The proportion judged grammatical for the unambiguous reparandum items was not as high as for the fluent unambiguous items, and this may have been partially due to an overall tendency for disfluent sentences to be judged less grammatical (marginally significant by subjects but not items).[6] However, the difference between the disfluent conditions was still relatively small, and it was unclear whether this was because information on the disfluency is a relatively weak influence on the following syntactic structure, or rather because ratings in the 90% range approach a ceiling value, making the range of variation small.

To address these concerns, we designed a second experiment in which we used the same kind of disfluencies in sentences that were the complement of those in the first experiment: Sentences with ambiguous verbs like *selected* were disambiguated to the preferred MV reading rather than the less preferred RR reading. Now the unambiguously past participle reparandum verbs are actually inconsistent with the final version of the sentence. In contrast, the structurally ambiguous reparandum verbs are, like the replacement verb, still biased to the MV reading and thus consistent with the final version of the sentence. If the structural information on the reparandum verb is a strong cue for upcoming structure, the condition in which the reparandum verb is inconsistent should cause more processing difficulty than the condition in which it is consistent. This should result in a larger effect than in the first experiment, in which both disfluency conditions were consistent. However, if the structural information on the reparandum verb is used only as a weak cue, we expect little or no difference between the two disfluency conditions, because they are both disambiguated to the MV reading that the parser has been shown to prefer.

EXPERIMENT 2

Method

Participants. Twenty-five undergraduates from Michigan State University participated in the experiment in exchange for course credit. All participants were native speakers of English and had normal hearing.

Materials. The stimuli consisted of 24 experimental items (Appendix B), most of which were modified items from Experiment 1. Each experimental item appeared in one of four variations of a sentence in which the matrix-verb reading was correct, rather than the reduced-relative reading (2). In this experiment, in the garden path conditions (2b–d), the sentence was completely disambiguated by the end of the NP

---

[6] A reviewer suggests that the main effect of disfluency may be due to the lack of obvious message-level motivation for the revisions in our disfluent materials (e.g., no striking difference in meaning between *chosen* and *selected*), leading them to be perceived as less natural.

(e.g., *the right answer*) although even the article (*the*) would in general be a sufficient cue for disambiguation. Note that the commas in (2) are included for expository purposes; all stimuli were presented aurally.

(2)

a. The little girl chose the right answer, so her teacher gave her a prize.
(fluent, unambiguous)

b. The little girl selected the right answer, so her teacher gave her a prize.
(fluent, ambiguous)

c. The little girl chosen-uh selected the right answer, so her teacher gave her a prize.
(disfluent, unambiguous reparandum)

d. The little girl picked-uh selected the right answer, so her teacher gave her a prize.
(disfluent, ambiguous reparandum)

In the fluent, unambiguous condition (2a), the reparandum verb was of an unambiguously simple past form, making the MV reading the only one possible. In the fluent, ambiguous condition (2b), the reparandum verb was ambiguous between a transitive or a past-participle. In the disfluent, unambiguous reparandum condition (2c), a verb replacement occurred in which the ambiguous replacement verb was preceded by a reparandum verb with an unambiguously past-participle form, like *chosen*. This reparandum verb would be inconsistent with the continuation of the sentence if not replaced by an ambiguous verb like *selected*. In the disfluent, ambiguous reparandum condition (2d), a verb replacement occurred in which the ambiguous verb was preceded by a reparandum verb that was similarly ambiguous between the two structures. On its statistically preferred reading, the reparandum verb would be consistent with the continuation of the sentence. We used an ambiguous, MV-preferred verb like *picked* in this condition rather than an unambiguously MV (simple past) verb like *chose* because we wanted the difference between conditions (2c) and (2d) as much as possible to represent the cost of having an inconsistent disfluency, rather than the benefit of having a completely disambiguating disfluency. The two fluent conditions were included as a baseline to show that the preference for the MV-reading in the ambiguous condition (2b) is strong enough that it is treated no differently than a sentence with an unambiguously MV verb.

It is possible to question whether replacements like (2c), in which both the actual verb and the thematic role assigned to the subject NP are changed, ever occur in real life. As we emphasised earlier, we do not claim that participants frequently have to deal with the ‘tricky’ kinds of verb replacements in our materials, but rather we examine the processing of such tricky replacements as a means of understanding processing of replacements and revisions in general, and as an alternate tool for studying the critical role of verb argument structures in language processing.

More specifically, it could be suggested that in Experiment 2, participants' sensitivity to 'strange' disfluencies could be the cause of any observed differences in processing difficulty. However, we would argue that this alternative explanation is implausible. Since participants were explicitly directed to consider only the corrected form of the sentence in their response, it seems unlikely that they would use their response to signal a *high-level* judgement about the felicity of the disfluency. If, on the other hand, the 'sensitivity' is thought of as an online, *implicit* recognition of the mismatch in the replaced thematic/lexical structure, it is essentially the same as the 'lingering inconsistent information' account that we proposed above.

Procedure. The procedure was the same as for Experiment 1.

Results

The grammaticality judgement data from Experiment 2 are presented in Figure 3. Data from one participant were not included because the participant fell below the 90% accuracy cutoff on judgements for the grammatical fillers. All analyses were done on arcsine transformed scores. In contrast to Experiment 1, the verbs used in the disfluency conditions did not map on to the verbs used in the non-disfluency conditions (e.g., fluent unambiguous verb = *chose*, disfluent unambiguous verb = *chosen*); therefore, doing a 2×2 ANOVA on these data as we did on the data collected in Experiment 1 did not seem appropriate. Instead, we limited our analysis to two planned comparisons on the arcsine transformed scores, one within the baseline fluent conditions and one within the disfluent conditions, which mirrored those planned comparisons done in Experiment 1.

Per cent judged grammatical for the grammatical and ungrammatical filler conditions were again in the expected range, at 97% and 18%, respectively. A paired-sample *t*-test on the unambiguous and the ambiguous fluent conditions showed no significant difference in proportion judged grammatical; by subjects, $t(23) = 0.483$, $p = .634$; by items, $t(23) = 0.440$, $p = .67$, as would be expected if the bias on the ambiguous verb towards a MV continuation were as strong as that of an unambiguously MV verb. However, a paired-sample *t*-test on the two disfluency conditions showed a significant difference in proportion judged grammatical; by subjects, $t(23) = 3.441$, $p < .01$; by items, $t(23) = 3.665$, $p < .01$, such that the unambiguous reparandum condition was judged grammatical significantly less often than the ambiguous reparandum condition.

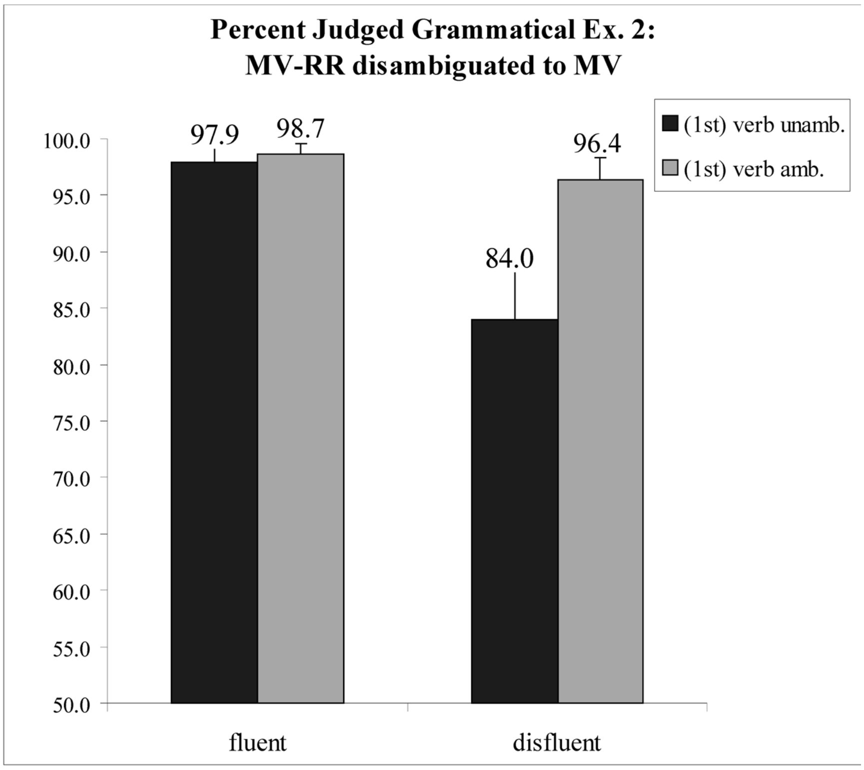

Figure 3. Percentage judged grammatical over conditions described in (2). Error bars represent standard errors.

Discussion

The results of Experiment 2 demonstrate that the structural information present in a replacement disfluency can have a strong influence on the parsing of the rest of the sentence: Sentences in which the structural information in the reparandum verb was inconsistent with the continuation were judged grammatical 13% less often than sentences in which the information was consistent, even though the failure to find a difference between the two fluent conditions indicated that the MV continuation in the ambiguous sentences by itself caused no processing difficulty compared with the unambiguous control. The fact that we found such a large effect of disfluency in this experiment is consistent with our speculation that the small size of the effect in Experiment 1 resulted from a ceiling effect rather than from an intrinsically small influence of such disfluencies on parsing.

We note that this experiment is the most successful attempt to date to test a strong prediction made by constraint-based models. Because constraint-based models hold that broad parsing preferences like Minimal Attachment and Late Closure are constraints on an equal footing with constraints over other types (semantic, pragmatic) and levels (lexical, discourse) of representations, they implicitly predict that in the presence of strong enough opposing constraints, the MV analysis should actually become more difficult. However, to our knowledge, an effect of this type has until now never been shown with the MV/RR garden path construction (although see Steinhauer, Alter, and Friederici 1999; Steinhauer and Friederici, 2001 for evidence that misplaced prosodic boundaries can give rise to a reverse garden path effect in the subordinate-main garden path). We found that the form of the reparandum verb in verb replacements does have such a degrading effect on the acceptability of the MV analysis, but we probably do not want to consider the verb form as a constraint in the traditional sense. Rather, we believe the form of the reparandum verb ''constrains'' the final analysis as a direct result of the specific mechanism by which the parser deals with repairs in the course of a sentence; this mechanism will be discussed in more detail below.

One caveat remains for the interpretation of the Experiment 2 results. The higher proportion of ungrammatical judgements in the unambiguous reparandum condition might be interpreted as a result of participants mis-remembering the stimulus. On this view, participants' responses did not reflect their sense of difficulty after their first parse, but rather, a more intensive process where they retrieved the sentence they just heard from working memory and judged this retrieved version. In disfluency cases, the sentence is sometimes incorrectly retrieved with the reparandum verb inserted as the correct verb, so that for example in Experiment 2, the response is actually based on a 're-parse' composed of *The girl chosen a piece of candy* rather than the feeling of difficulty/parse failure engendered by *The girl chosen-uh selected a piece of candy*. In fact, taking this idea further, one could suppose that the facilitation of garden path processing with unambiguous verbs in Experiment 1 was due to some number of trials in which the sentence was retrieved and re-parsed with the unambiguous verb, eliminating any possibility of a garden path. If this mis-retrieval account were correct, the results of the experiment would show that the processing, storage, and retrieval of sentences with disfluencies are directly influenced by the content of the disfluency, but that these effects are not realised immediately during the parse.

It is somewhat of a challenge to test this possibility directly since the preferred online measures typically used in sentence comprehension experiments do not lend themselves to auditory presentation (e.g., eyetracking during reading). It would also be hard to generalise from performance on an explicit sentence recall task (e.g., Potter & Lombardi, 1998) to the 'unconscious' recall that this view says takes place at the judgement point. In addition, it seems fairly unlikely to us that, in an experiment where a judgement prompt was expected at the end of each sentence, participants would often retrieve and re-parse the sentence in order to respond. (Notice that this retrieval would have to happen rather frequently across trials in order to give rise to a mis-retrieval rate of 16%.) Therefore, although with the limited real-time precision of our current auditory comprehension measures we cannot completely rule out a version of our story where disfluency affects later processes, we consider the online explanation to be the more parsimonious one, and we have structured our discussion accordingly. And with a global sense of what the patterns of performance look like, we intend to conduct experiments in the future using techniques that allow moment-to-moment measurement of processing difficulty (e.g., eyetracking while listening to spoken instructions that contain disfluencies). These studies will permit more detailed analyses of our conclusions.

GENERAL DISCUSSION

The results from Experiments 1 and 2 show that structural information associated with the reparandum verb in a verb replacement disfluency can influence processing of the rest of the sentence. The findings thus add another piece to the growing body of evidence that disfluencies affect comprehension in complex ways, and are not simply filtered out by some form of pre-processing. But more specifically, these results begin to address questions about the mechanisms by which listeners recover from ‘‘lexically-rich’’ disfluencies like revisions, and potentially, issues about the mechanisms by which listeners deal with complex fluent sentences as well.

Mechanisms underlying verb replacement effects

We have shown that the argument structure of the reparandum verb in a verb replacement disfluency can influence the parsing of the rest of the sentence. But by what mechanism does this influence make itself felt? One possibility we have mentioned is that this observed influence is a result of priming of the potential syntactic frames associated with verbs. Such an effect has been shown by Trueswell and Kim (1998) in a series of fast-priming experiments. Since their fast-priming design resembles our disfluency manipulation, it is worth examining their study in more detail.

Trueswell and Kim (1998) gave participants a self-paced reading task including sentences exhibiting the direct object/sentence complement (DO/SC) ambiguity, in which the DO-biased main verb was sometimes ‘‘fast-primed’’ by first flashing a different verb for an extremely brief amount of time (39 ms), such that participants were conscious only of a ‘‘flicker’’. Verb primes with a SC-bias significantly diminished and verb primes with a DO-bias significantly increased the strength of the garden path (as measured by reading times at the disambiguation point), where both cases were compared to equibiased primes. They argued from these results that recognition of the prime verb activated ''invariant'' syntactic and semantic features (which we can also think of as abstract elementary trees), which then facilitated the selection of the same features for the real main verb. The design is thus similar to our Experiment 1, in which participants encountered an RR-''biased'' verb as part of a disfluency before encountering the MV-biased replacement verb. This similarity highlights the possibility that the effects that we found are not due strictly to disfluency-specific mechanisms, but are rather a result of general priming effects that can be observed whenever a verb and its associated lexical information are presented during a sentence.

It seems only sensible to assume that a priming mechanism like that demonstrated in Trueswell and Kim's (1998) study plays some role in the effects we found with disfluencies, because both experiments involved presentation of an ''experimental'' verb in the close vicinity of the ''real''/replacement verb. However, there are several reasons (over and above the obvious differences in presentation modality) why we believe that the disfluency findings are probably not completely reducible to the priming effects Trueswell and Kim observed. One issue is that of the participants' consciousness of the potentially priming material. Trueswell and Kim argue that their results are meaningful in part because most participants did not consciously perceive the verb primes. In fact, based on a mini-analysis of the few participants who did report awareness of the primes, they speculated that this awareness might diminish the effect by triggering inhibition of the prime. However, participants in our experiments who heard a verb replacement disfluency were certainly aware of the reparandum verb, even if (as the Lickley and Bard (1996) findings might suggest) this awareness is fleeting. Simply put, the comprehender must have some temporary realisation that a revision has occurred and that the reparandum verb was a mistake in order to build an acceptable structure. However, our findings here show that in these cases the effect of the structural information still holds, and thus has not been completely inhibited. In fact, as we discussed earlier, versions of an inhibition account which assume that lexically associated structures can be independently activated would predict that processing the replacement in an RR-structure would actually be *harder* in cases where the reparandum verb has this structure associated with it, which of course is the opposite of what we see here. Therefore, if Trueswell and Kim are right in suggesting that subliminal presentation is key to the verb priming effect they found, it looks like the disfluency effect must in some way be different.

A related and more central point is that Trueswell and Kim's (1998) priming mechanism manipulates lexical representations (and their associated structures) only. In other words, the prime verb's influence is directly felt only on the choice between possible syntactic frame representations localised on the sentence's ''true'' verb, and it is only indirectly, through the choice of one of these frames, that a global analysis of the sentence is affected. However, under the assumption that parsing is incremental, revisions are importantly different from masked presentation of verbs: In revisions, the ''prime'' verb is actually inserted/ substituted into an early, partially-completed sentence representation (or into more than one, in parallel models), and in this structure the verb's arguments begin to be filled in with NPs so far encountered. In our experiments, therefore, s*ome* kind of direct manipulation of global structure is necessary when the correct verb is encountered, whether it be completely deleting the existing structure to start over, or something more complex.

Work we have recently done on verb replacements that involve a different structural alternation has supported the idea that lexical priming is not the only force responsible for these effects. In Ferreira, Lau, and Bailey (2004), we examined verb replacements in which the disfluency verb and the replacement verb differed in whether a third argument was optional (*drop* the ball [in the box]) or obligatory (*put* the ball in the box). We found that a disfluency like *Simon says to put, uh, drop the ball* caused the proportion of sentences judged grammatical to drop compared with a fluent control condition, presumably because the *put* disfluency in some way created an expectation for a third argument that was not fulfilled. This finding could potentially be explained with lexical priming, if the obligatory 3-argument tree of *put* primed the 3-argument tree of *drop*. However, we also found the reverse: A disfluency in an ungrammatical sentence like *Simon says to drop, uh, put the ball* actually caused the proportion judged grammatical to rise, presumably because the *drop* disfluency somehow diminished the expectation for a third argument. This result is difficult to explain with lexical priming, because *put* should not have an associated 2-argument tree available for priming. To explain the *put-drop* results as well as the experiments described here, the proposed mechanism by which the processor deals with disfluencies must directly operate over representation(s) created at the sentence level.

Implications for Brennan and Schober (2001)

Our experiments were partly motivated by Brennan and Schober's (2001) study, which suggested that the cost of processing revisions varies depending on the degree to which the reparandum information is entrenched in memory and the length of time available for inhibiting/

deleting this information. We hypothesised that if the length of time available was insufficient for the processor to delete structural information present in the replacement or to reverse parsing decisions made based on this false information, this structural information might continue to exert effects on comprehension. We did find such an influence of structural information, consistent with Brennan and Schober's work. However, our results do not show whether this influence is a direct byproduct of time and resource limitations. It might be that the parser does not even attempt to inhibit the structural information, but in fact uses information from the reparandum to predict how to parse the repair. Whether it does this passively, through failure to inhibit activation of reparandum material, or actively, using the reparandum material as one of many constraints on the parsing of the repair, is still a question.

Overlay

We have recently developed a way of characterising the process within an LTAG framework by introducing an operation called *Overlay* (Ferreira et al., 2004). The implications of Overlay are worked out more thoroughly in that paper; here we will discuss the operation with reference to the present set of results. We assume the simplest version of an LTAG parser that, upon encountering a lexical item, retrieves an associated elementary tree, looks for a substitution site in the current structure, and inserts the elementary tree into that structure.[7] Overlay comes into play only when the parser retrieves an elementary tree and cannot find an open substitution site, as may happen with disfluencies like repetitions or revisions.[8] When this happens, the parser uses root node identities to find the appropriate part of the existing structure onto which it can overlay the recently retrieved tree. The resulting representation will have some features of the multi-dimensional phrase structures assumed by Williams (1978) for phrasal coordination (and more recently by Goodall (1987) and Moltmann (1992), among others). In the case of repetition, the overlaying trees will be identical, but in the case of revision, they will differ at certain points. We assume that the most recent tree will ‘‘win’’ in the determination of the final structure—if

[7] TAGs incorporate two fundamental operations, substitution and adjunction. At this point, we limit our hypothesised Overlay operation to substitution cases, although the implications of Overlay for adjunction is certainly an interesting issue which we plan to consider in the future.

[8] It is possible that, at least in some cases, the signal to the parser to try to initiate Overlay is elicited by the failure of the incoming word to match a structural prediction (as in prediction of a preposition in *The little girl chosen* . . .) rather than simple failure to find a substitution site.

only because it has been more recently activated—but that the previous structure is not completely erased and thus can still exert effects on parsing like the ones we find. Overlay, then, can account for the *put-drop* experiments which could not be explained by lexical priming.

In the verb-replacement experiments described in this paper, much of the influence of the reparandum verb probably results from lexical priming which in Experiment 1, for example, makes the selection of the RR elementary tree associated with the repair verb more likely. The fact that the actual magnitude of the effect in this study was considerably larger than in the *put-drop* study is consistent with the idea that the *put-drop* effect was mainly due to Overlay and the *chosen-selected* effect was due to Overlay + lexical priming. Another factor that may influence the initial elementary tree selection here is that the first NP was initially assigned the theme role on the basis of the reparandum verb; if one constraint on selection of verb argument structure is NP animacy (as argued in Trueswell, Tanenhaus, & Garnsey, 1994), then it may well be that some kind of *theme*-marking on the first NP could also act as a constraint here. If the RR tree is chosen for these reasons, then the Overlay operation will obviously be easy; the activated *selected* RR tree can simply be placed ''on top of'' the previous *chosen* tree. We can see how this strategy could cause the parser to fail by looking at Experiment 2: Here the ultimately incorrect RR elementary tree will be primed by the reparandum verb, and whichever tree is actually selected will be overlaid on the already-built structure, which will be incompatible with the continuation. Thus, on a certain number of trials the underlying structure will retain enough activation to disrupt the parse. When it works, however—which may be most of the time in real life—Overlay makes it possible for the partially-built, global sentence representation begun earlier to be conserved, and thus avoids unnecessary and costly deletion of structure.

Many questions remain to be answered about the way Overlay operates. One issue is that the two fundamental structures that participate in Overlay—the original, which ends up at the ''bottom'', and the correct tree, which ends up on ''top''—will not always have completely reconcilable shapes. For example, the original tree might be anchored by a complex verb (*has been put*) rather than a simple past tense verb (*put*), and as a result the verbs' postverbal arguments will be at different heights in the bottom and top structures. How Overlay then operates and influences interpretations becomes an important question for future work, although we suspect that simply because the top tree has priority and the bottom tree is supposed to be in some sense abandoned, the parser will simply Overlay the two trees at the identical root nodes and give up on the original if its shape is radically different from the corrected structure.

Another interesting question is how we should integrate the Overlay operation with the role that prosody has been shown to play in disfluency processing. It might be that certain prosodic phenomena act as signals that the parser should implement Overlay, and intonational cues may guide the process of aligning the two trees. (Indeed, it is interesting to note in this regard that Selkirk (1996) argues for intonational constraints on the right-node-raising constructions for which multidimensional phrase structures have often been proposed.) Thus, the prosodic disruption associated with the repair might help to inform the parser that the input contains a disfluency. In addition, because the regular intonational structure must be resumed at the repair, it is possible that the prosodic system engages in an intonational analogue of the syntactic Overlay operation we have proposed. The two forms of Overlay could then work together to help the parser locate the appropriate site for reconciling the reparandum and repair structures.

Because until recently psycholinguistic studies of disfluency processing have been fairly sparse, more experimental work is needed to determine whether the Overlay operation we propose is really the optimal way to describe the relevant mechanism. We acknowledge that the proposal requires a great deal more elaboration, but we believe it is an important first step because it captures the fundamental ideas that any theory of disfluency reanalysis will have to explain. These are that at some level the parser must (1) recognize it has encountered a disfluency involving a reparandum and a repair, (2) find the point at which the reparandum began, and (3) get rid of the reparandum material (words and structure) from the ongoing parse and replace it with the corrected material. Our proposal concerning Overlay captures all three of these features: (1) the parser recognises that it has a disfluency on its hands when it finds itself with two trees that cannot be unified through regular TAG operations; (2) the parser locates the point where the reparandum began by finding the root node in the first tree that is identical to the root node of the new tree; and (3) the reparandum is replaced when the new tree is placed on top of the old one, with the two trees anchored at the identical root nodes. Moreover, because Overlay is based on the formal machinery of LTAG, it is possible to imagine how the processes of disfluency reanalysis can eventually be implemented in a formal model, allowing additional predictions to be rigorously derived and tested.

Disfluencies and garden path reanalysis

Our speculation that lingering disfluency material might influence sentence comprehension was inspired by recent studies that, while not specifically referencing disfluencies, have demonstrated effects of lingering early representations on later processing. For example, Christianson et al. (2001) have shown that the initial misparse of a garden path sentence like *While Anna dressed the baby played in the crib* has lingering effects on the semantic representation built (see also Ferreira, Christianson, & Hollingworth, 2001). Their data show that even when the initial misinterpretation of a garden path sentence seems to be reanalysed—as indicated by accurate responses to comprehension questions probing the main clause—the initial misinterpretation ''lingers'' and results in inaccurate answers to questions probing the earlier, subordinate clause. For example, participants almost always answered ''yes'' to questions like *Did the baby play in the crib?,* indicating that the parser had successfully recovered from the initial analysis of *baby* as object of the subordinate clause so that it could reassign *baby* as the subject of the matrix clause. However, participants also largely answered ''yes'' to questions like *Did Anna dress the baby?,* suggesting that the initial interpretation was still somewhat active. The Christianson et al. work supports the idea that the reanalysis process is in some respects conservative. When the parser encounters a problem, it engages only in the minimal steps necessary to re-attain a ''good-enough'' threshold of acceptability; it does not bother to exhaustively clean up the pieces from the old analysis. As we have pointed out, the process by which the parser must recover from a revision is reminiscent of reanalysis—a commitment is made which turns out to be incorrect—and it seems that, as in the garden path case, the parser does not thoroughly delete the original (reparandum) material and structure associated with it.

The similarity between these findings suggests to us that the process of ''recovering'' from revisions has important parallels to processes of reanalysis. In both cases the parser has to go back and reverse earlier decisions—whether these decisions are seen as structure-*building* (serial models) or structure-*selecting* (parallel models). Although at such an early stage of research on disfluencies it is hard to judge the strength and utility of this isomorphism, we think it likely that, with more study, disfluency may turn out to be a useful phenomenon for approaching questions of ''regular'' syntactic reanalysis. The disfluency processing mechanisms proposed here are tantalisingly reminiscent of models of reanalysis in which the repair process is made up of a series of minimal steps that continue only until the parse reaches some threshold level of acceptability (e.g., Fodor & Inoue, 1998). Our experiments suggest that in dealing with revisions, the parser analogously makes the minimal changes necessary to accommodate the replacement of material without necessarily ''cleaning up'' all of the other syntactic/thematic consequences of the change. Of course, such links at this stage remain highly speculative.

CONCLUSIONS

In these experiments we have shown that structural information encoded in disfluencies like revisions can influence sentence processing. Many questions about the mechanism and extent of this influence remain to be answered, but our work at the least indicates a new avenue of research in what is still a surprisingly understudied area of sentence comprehension. Although assuming an idealised linguistic input allows psycholinguists to address many interesting questions, we often forget that many of the ‘‘ecological’’ issues in language comprehension that are thus glossed over are real, frequent, and quite interesting in their own right. In the present case the problem that arises when, in the midst of the construction of a sentence representation input is retracted, seems hardly less challenging to the comprehension system or worthy of interest to psycholinguists than the extensively studied problem of dealing with input that is temporarily ambiguous.

With our experiments on verb replacement, we have begun to shed light on both the specifics of this problem and some characteristics of the strategy that the parser uses to deal with it. In addition, our overall body of work on disfluency has given empirical support to theories of basic comprehension processes such as the Head Position Effect and the importance of lexical information for sentence processing. We therefore believe our work to be a good demonstration of the ways in which studying disfluency can be informative and worthwhile on the large scale as well as in its own domain.

Manuscript received August 2003
Revised manuscript received May 2004

REFERENCES

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

Bader, M. (1998). Prosodic influences on reading syntactically ambiguous sentences. In J. D. Fodor and F. Ferreira (Eds.), *Reanalysis in sentence processing* (pp. 1–46). Dordrecht, the Netherlands: Kluwer Academic.

Bailey, K. G. D., & Ferreira, F. (2003). 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.), *Approaches to studying world-situated language use: Bridging the language-as-product and language-as-action traditions*.

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.

Charniak, E., & Johnson, M. (2001). Edit detection and parsing for transcribed speech. In *Proceedings of the 2nd Meeting of the North Amercian Chapter of the Association for Computational Linguistics*, pp. 118–126. Pittsburgh, PA.

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.

Core, M., & Schubert, L. (1999). *Speech repairs: A parsing perspective*. Paper presented at the International Congress of Phonetic Sciences, Satellite Meeting on Disfluency in Spontaneous Speech, University of California, Berkeley, CA, July.

Ferreira, F. (2000). Syntax in language production: An approach using tree-adjoining grammars. In L. Wheeldon (Ed.), *Aspects of language production*. Hove, UK: Psychology Press.

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. M. (1991a). Recovery from misanalyses of garden path sentences. *Journal of Memory and Language*, 25, 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). Amsterdam: North-Holland.

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

Ferreira, F., Lau, E. F., & Bailey, K. G. D. (in press). Disfluencies, language comprehension, and tree adjoining grammars. *Cognitive Science*.

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

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.

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.

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

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

Gibson, E. (1991). *A computational theory of human linguistic processing: Memory limitations and processing breakdown*. Ph.D. thesis, Carnegie Mellon University, May.

Goodall, G. (1987). *Parallel structures in syntax*. Cambridge: Cambridge University 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.

Kim, A., Srinivas, B., & Trueswell, J. C. (1999). The convergence of lexicalist perspectives in psycholinguistics and computational linguistics. In P. Merlo & S. Stevenson (Eds.), *Sentence processing and the lexicon: Formal, computational and experimental perspectives*. Philadelphia, PA: John Benjamins Publishing.

Lickley, R. J. (1995) Missing disfluencies. In *Proceedings of International Congress of Phonetic Sciences*, *4*, 192–195. Stockholm.

Lickley, R. J. & Bard, E. G. (1992). Processing disfluent speech: How and when are disfluencies found? In *Proceedings of the International Conference on Spoken Language Processing*, 935–938. Banff, October.

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, PA.

Lickley, R. J., Shillcock, R. C., & Bard, E. G. (1991). Processing disfluent speech: How and when are disfluencies found? In *Proceedings of the Second European Conference on Speech Communication and Technology, Vol. III,* 1499–1502. Genova, September.

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.

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

Moltmann, F. (1992). *Coordination and comparatives*. Doctoral dissertation, MIT, Cambridge, MA, USA.

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.

Potter, M. C., & Lombardi, L. (1998). Syntactic priming in immediate recall of sentences. *Journal of Memory and Language*, *38*, 265–282.

Pritchett, B. (1988). Garden path phenomena and the grammatical basis of language processing. *Language*, 64(3), 539–576, 1988.

Selkirk, L. (1996). *Constraints on prosodic phrasing and their relevance for right node raising*. Talk presented at MIT, Linguistics Colloquium, December.

Shapiro, L. P., Zurif, E., & Grimshaw, J. (1989). Verb processing during sentence comprehension: Contextual impenetrability. *Journal of Psycholinguistic Research*, *18*, 223–243.

Shriberg, E. E. (1996). Disfluences in SWITCHBOARD. *Proceedings of the International Conference on Spoken Language Processing*. [Addendum, 11–14], Philadelphia, PA, USA.

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

Steinhauer, K. & Friederici, A. D. (2001). Prosodic boundaries, comma rules, and brain responses: The closure positive shift in ERPs as a universal marker for prosodic phrasing in listeners and readers. *Journal of Psycholinguistic Research*, *30*, 267–295.

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

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

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

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

APPENDIX A

1. The famous athlete driven / escorted / driven-uh escorted / accompanied-uh escorted to the arena admired the brand-new facilities.
2. The Russian diplomat driven / accompanied / driven-uh accompanied / escorted-uh accompanied to the White House talked heatedly with the president.
3. The ruthless dictator overthrown / seized / overthrown-uh seized / captured-uh seized in the coup fled to a neighbouring country.
4. The corrupt rulers overthrown / captured / overthrown-uh captured / seized-uh captured during the war never staged a comeback.
5. The thoroughbred horses shown / raced / shown-uh raced / exhibited-uh raced at the fair wore black and red ribbons.
6. The antique cars shown / exhibited / shown-uh exhibited / raced-uh exhibited at the auto exposition were valued at millions of dollars.
7. The little girl chosen / selected / chosen-uh selected / picked-uh selected for the role celebrated with her parents and friends.
8. The job applicant chosen / picked / chosen-uh picked / selected-uh picked for the position received an unbelievable salary.
9. The suspected muggers seen / recognised / seen-uh recognised / discovered-uh recognised in their car gave no evidence of guilt.
10. The rebellious kids seen / discovered / seen-uh discovered / recognised-uh discovered behind the fence already skipped six classes.
11. The tiny insects eaten / consumed / eaten-uh consumed / devoured-uh consumed in the night are the bats' major food source.
12. The Arctic caribou eaten / devoured / eaten-uh devoured / consumed-uh devoured during the winter saved the wolves from starvation.
13. The nude model drawn / used / drawn-uh used / depicted-uh used in art class stood completely still for hours.
14. The handsome actor drawn / depicted / drawn-uh depicted / used-uh depicted in the publicity brochure caught the director's attention.
15. The tall man beaten / hit / beaten-uh hit / punched-uh hit in the riot was taken to the hospital.
16. The gang member beaten / punched / beaten-uh punched / hit-uh punched during the street fight vowed to get revenge.
17. The young woman shaken / jolted / shaken-uh jolted / tossed-uh jolted during the subway ride stumbled onto the platform.
18. The weary travellers shaken / tossed / shaken-uh tossed / jolted-uh tossed many times during the journey felt nauseous.
19. The student group taken / led / taken-uh led / guided-uh led through the museum had lots of questions.
20. The police officers taken / guided / taken-uh guided / led-uh guided through the company's warehouses were not impressed.
21. The elderly neighbour bitten / assaulted / bitten-uh assaulted / attacked-uh assaulted by the Rottweiler filed a complaint.
22. The shark victim bitten / attacked / bitten-uh attacked / assaulted-uh attacked the most violently sustained severe injuries.
23. The marathon runner overtaken / defeated / overtaken-uh defeated / passed-uh defeated at the finish line looked very upset.
24. The racecar driver overtaken / passed / overtaken-uh passed / defeated-uh passed in the last lap demanded a rematch.

APPENDIX B

1. The famous athlete drove/ escorted / driven-uh escorted / accompanied-uh escorted his girlfriend to the arena in a limousine.
2. The Russian diplomat drove/ accompanied/ driven-uh accompanied/ escorted-uh accompanied the African ambassador to the White House for the party.
3. The ruthless dictator overthrew/ seized/ overthrown-uh seized/ captured-uh seized the opposition in a bloody coup supported by the military.
4. The frustrated rebels overthrew/ captured/ overthrown-uh captured/ seized-uh captured the government leaders during the civil war.
5. The little girl showed/ raced/ shown-uh raced/ chased-uh raced her favorite cousin to the guest room in the attic.
6. The high school kids showed/ exhibited/ shown-uh exhibited/ displayed-uh exhibited very little interest in the educational video.
7. The manager chose/ picked/ chosen-uh picked/ selected-uh picked the job applicant who had dressed the most fashionably.
8. The little girl chose/ selected/ chosen-uh selected/ picked-uh selected the right answer, so the teacher gave her a small prize.
9. The suspected robbers saw/ recognised/ seen-uh recognised/ discovered-uh recognised the cashier who took the witness stand.
10. The rebellious kids saw/ discovered/ seen-uh discovered/ recognised-uh discovered a secret tunnel that led underneath the school.
11. The bats ate/ consumed/ eaten-uh consumed/ devoured-uh consumed millions of tiny insects in the hours before the sun came up.
12. The wolves ate/ devoured/ eaten-uh devoured/ consumed-uh devoured the Arctic caribou that they had killed hours before.
13. The bedridden artist drew/ depicted/ drawn-uh depicted/ presented-uh depicted/ the peaceful meadows right outside his window.
14. The art student drew/ presented/ drawn-uh presented/ depicted-uh presented the wealthy woman in a flattering classical pose.
15. The vandals beat/ hit/ beaten-uh hit/ punched-uh hit people with sticks and clubs as the riot continued.
16. The gang member beat/ punched/ beaten-uh punched/ hit-uh punched his rival during the street fight.
17. The old train shook/ jolted/ shaken-uh jolted/ tossed-uh jolted the passengers as it rattled down the track.
18. The little girl shook/ tossed/ shaken-uh tossed/ jolted-uh tossed the puppy in the air playfully.
19. The guard took/ guided/ taken-uh guided/ led-uh guided the student group on a tour of the old cemetery.
20. The night watchman took/ led/ taken-uh led/ guided-uh led the police officers through the company's warehouses.
21. The couple's Rottweiler bit/ assaulted/ bitten-uh assaulted/ attacked-uh assaulted the old man as he fumbled with his keys.
22. The shark bit/ attacked/ bitten-uh attacked/ assaulted-uh attacked the surfer that had gone too far from shore.
23. The marathon runner overtook/ defeated/ overtaken-uh defeated/ passed-uh defeated the others just before the finish line.
24. The racecar driver overtook/ passed/ overtaken-uh passed/ defeated-uh passed all of his competitors after he changed his tyres.