Differences in the Timing of Implausibility Detection for Recipient and Instrument Prepositional Phrases

Allison Blodgett[1,3] and Julie E. Boland[2]

We conducted two word-by-word reading experiments to investigate the timing of implausibility detection for recipient and instrument prepositional phrases (PPs). These PPs differ in thematic role, relative frequency, and possibly in argument status. The results showed a difference in the timing of garden path effects such that the detection of implausible dative recipients (which are clearly arguments) was delayed relative to the detection of implausible instruments (which may not be arguments). They also demonstrated that commitments to syntactic structure were made at the preposition for both dative and instrument PPs. While these results refute delay models of parsing (e.g., Britt, 1994) and syntax-first accounts of PP-attachment (e.g., Frazier, 1978; Frazier & Clifton, 1996), they support constraint-based lexicalist models that enable verb bias and plausibility information to compete (Garnsey, Pearlmutter, Myers, & Lotocky, 1997).

KEY WORDS: Parsing; plausibility; prepositional phrase attachment; sentence comprehension; syntactic ambiguity.

INTRODUCTION

Prepositional phrase (PP) attachment has been widely studied among those investigating the nature of the human sentence processing mechanism. For example, PPs have been used to investigate the role of minimal attachment (Clifton, Speer, &

[1] Department of Linguistics, The Ohio State University, Columbus, Ohio 43210.
[2] Department of Psychology, University of Michigan, Ann Arbor, MI 48109.
[3] To whom all correspondence should be addressed: OSU Department of Linguistics, 1712 Neil Avenue, #222, Columbus, OH 43210. email: blodgett@ling.ohio-state.edu

Abney, 1991; Rayner, Carlson, & Frazier, 1983), thematic role information (Taraban & McClelland, 1988), argument/adjunct status (Britt, 1994; Schutze & Gibson, 1999; Speer & Clifton, 1998), and frequency information (Boland & Boehm-Jernigan, 1998; Spivey-Knowlton & Sedivy, 1995) on the resolution of syntactic ambiguity. Yet in spite of this work, a consensus regarding the parsing of PPs has not emerged. There is still disagreement as to whether attachment decisions are made at the preposition (Boland & Boehm-Jernigan, 1998) or postponed until the entire PP constituent has been constructed (Britt, 1994; Britt, Gabrys, & Perfetti, 1993). There is disagreement as to whether attachment decisions are made according to the principles of minimal attachment and late closure (Frazier, 1978, 1990; Frazier & Clifton, 1996; Traxler, Pickering, & Clifton, 1998) or via the convergence of multiple, weighted constraints (Boland & Boehm-Jernigan, 1998; Taraban & McClelland, 1988). Additionally, proponents of constraint-based models disagree as to whether all constraints are used simultaneously (e.g., MacDonald, Pearlmutter, & Seidenberg, 1994; Trueswell, Tanenhaus, & Garnsey, 1994) or whether there is a distinction between constraints that affect the generation of one or more syntactic structures and constraints that affect the selection of structure (Boland, 1997; Boland & Boehm-Jernigan, 1998; Lewis & Boland, 2000).

In this paper, we focus on one aspect of PP-attachment—the timing of the detection of an implausible prepositional object—because we believe that differences in timing can help distinguish among three approaches to PP-attachment: syntax-first models, constraint-based lexicalist models, and delay models.

PPs are often studied in sentences of the form <subject, transitive verb, direct object, PP> because the PP can modify either the verb phrase (VP) or the direct object as either an argument or an adjunct. As shown in the example stimuli in (1) and (2) below, the sentences in our materials were no exception. We used dative and action verbs to create three stimulus sets in which we manipulated the preposition in the first PP following the verb (underlined in the examples) and the plausibility of the embedded noun phrase (NP; marked in bold). For the two dative verb sets, the embedded NP was either a plausible or implausible recipient, and for the action verb set, the embedded NP was a plausible or implausible instrument. In each of the critical sentences, the underlined PP must be interpreted as modifying the immediately preceding NP in order for the sentence to make sense.

(1) Example of Dative PP Stimuli[4]
 a. John gave a letter <u>to</u> **his son** to a friend a month ago.
 b. John gave a letter <u>about</u> **his son** to a friend a month ago.
 c. Paul gave the script <u>to</u> **the play** to a girl during the break.
 d. Paul gave the script <u>for</u> **the play** to a girl during the break.

---

[4] We manipulated the plausibility of the embedded NP as a recipient (*his son, the play*) as a between-items variable in the dative PP stimuli.

(2) Example of Instrument PP Stimuli
 a. The repairman sealed a container <u>with **clear glue**</u> with a lid, while he smiled.
 b. The repairman sealed a container <u>of **clear glue**</u> with a lid, while he smiled.
 c. The repairman sealed a container <u>with **clear knobs**</u> with a lid, while he smiled.
 d. The repairman sealed a container <u>of **clear knobs**</u> with a lid, while he smiled.

In order to understand the competing predictions that these materials test, it is necessary to look more closely at the differences between dative and instrument PPs. Dative *to*-PPs are unquestionably arguments of the verb, and they are assigned the thematic role of recipient. As argument phrases, they depend in part on the words they modify for their semantic interpretation (Clifton *et al*., 1991; Schutze & Gibson, 1999), and they tend to behave in predictable ways syntactically (see the argumenthood tests of Schutze and Gibson). One customary way of representing this relationship in linguistic and psycholinguistic theories is to have lexical entries contain information about the syntactic category and semantic role of the potential arguments of a word. Thus, for the dative *to*-PPs used in our materials (adopted from Boland & Boehm-Jernigan, 1998), activating the dative verb during reading or listening might simultaneously activate information about the argument PP.

In contrast to dative PPs, there is some suggestion that instrument *with*-PPs are not arguments, but adjuncts. They fail two of the syntactic tests for argumenthood,[5] and because *with* assigns multiple thematic roles, as shown below in (3), the role of the PP cannot be assigned with any degree of confidence until the NP has been evaluated. Traditionally, adjunct phrases are not represented in the lexical entries of the words they modify, although this is not true of all linguistic and psycholinguistic accounts (e.g., Cinque, 1999; MacDonald *et al*., 1994).

(3) Examples of the Multiple Thematic Roles Assigned by *With*
 a. The repairman sealed the container with clear glue.   (instrument)
 b. The repairman sealed the container with his boss.   (co-agent)
 c. The repairman sealed the container with a smile.   (property of VP)
 d. The repairman sealed the container with a crack.   (property of NP)
 e. The repairman put the container with the shovel.   (location)
 f. The repairman loaded the container with the supplies.   (object/theme)

---

[5] For a discussion of instruments and argument tests see Schutze (1995) and Sedivy and Spivey-Knowlton (1994). In brief, instruments behave like adjuncts when they follow known adjuncts (e.g., *The repairman sealed a container on Tuesday with clear glue*) and when they occur with VP pro-forms (e.g., *The repairman sealed a container with clear glue, but his boss did so with caulking*).

In our dative and instrument stimuli, the plausibility of the first PP as the recipient or instrument of the verb is affected by two pieces of information: the preposition itself and the embedded noun phrase. For example, *to* typically marks a recipient role when it follows a dative verb. The prepositions *of* and *about*, however, cannot assign that role. Likewise, while *with* can mark an instrument role when it follows an action verb, *of* cannot. Similar properties hold for the embedded NP. While sons are likely recipients, plays are not; and while glue is a likely instrument, knobs are not. Note that if the underlined PP is mistakenly VP-attached in *John gave a letter to his son to his friend*, the error does not become apparent until *his friend*. This is because prior to *his friend*, the second *to* can be interpreted as an infinitive (e.g., *John gave a letter to his son to mail*) or as NP-attached (e.g., *A lawyer revealed the payment to the witness to the crime*). Likewise, the VP-attachment of the underlined PP in *The repairman sealed a container with clear glue with a lid* does not become infelicitous until *a lid* is encountered. The second *with* could be analyzed as heading either another VP adjunct (e.g., *with his brother*) or an NP adjunct (e.g., *with glitter on it*).

Although these materials are similar in that the plausibility manipulations were controlled along the same two dimensions, different theories make different predictions about when the implausibility of the embedded NP should become apparent at least when the preposition is consistent with assignment of a recipient or instrument thematic role, as in (1c) and (2c). Note that implausibility detection must logically be distinct from and prior to whatever reanalysis processes result.

Syntax-first models such as the traditional garden path model (e.g., Frazier, 1978, 1990) and construal (e.g., Frazier & Clifton, 1996; Traxler *et al.*, 1998) predict that both verb types will elicit the same pattern. PPs in a verb-NP-PP structure are initially treated as "primary phrases," and thus a decision is made at the preposition to attach the PP onto the VP as an argument according to minimal attachment. If we assume that dative *to*-PPs and instrument *with*-PPs are both arguments, then there should be no difference in the timing of the detection of an implausible embedded NP. Upon encountering the *to*-PP or *with*-PP, the recipient or instrument role of the verb would be assigned to the PP, and the thematic processor would immediately check the goodness-of-fit between the roles and the embedded NPs. If this view is correct, implausibility detection is likely to be visible right at the NP itself in (1c) and (2c) compared to (1a) and (2a).

Note that even if instrument *with*-PPs are adjuncts, syntax-first models would still not necessarily predict a difference in the timing of the detection of an implausible instrument NP relative to a recipient PP. Each PP in a verb-NP-PP structure would still be initially treated as a primary phrase, and the PP would still be attached to the VP as an argument at the preposition. The recipient role of the dative verb would be assigned to the *to*-PP, and the goodness-of-fit of the embedded NP would be evaluated as above. In contrast, the action verb would have no thematic role to assign to the *with*-PP. This should trigger reanalysis of the

with-PP as a VP-adjunct,[6] thereby revealing the implausibility of (2c) compared to (2a). Thus, implausibility detection could still occur at the embedded noun for both verb types, although there is the possibility that instruments will be delayed relative to datives because instrument PPs require an additional stage of reanalysis prior to implausibility detection.

In a slightly different scenario, imagine that the combination of the action verb and *with* triggers the thematic processor to construe the PP to the VP, the most recent theta domain in all of our materials. Under this account, there might be no cost associated with the implausible instrument PP because it can be interpreted as modifying the direct object NP contained within the theta domain. However, this alternative seems unlikely simply given native speaker intuitions that the critical *with*-PPs in these materials are initially interpreted as instruments. Another possible prediction of this construal account is increased processing load at the embedded NP, either because multiple attachment sites and interpretations are being considered in parallel or because the thematic processor detects the implausibility of the NP as an instrument of the verb during those trials in which that attachment is built first (Traxler *et al*., 1998). Regardless of which explanation one adheres to, none of the syntax-first scenarios predict that detection of an implausible instrument will happen earlier than the detection of an implausible recipient.

Our materials force traditional constraint-based lexicalist models to produce the same outcome as syntax-first models. This is because the disambiguating embedded NP occurs two words after the point of ambiguity, and an attachment decision would be made at the preposition, consistent with the information from multiple constraints. For both sets of materials, the PP would likely be VP-attached and assigned a thematic role. In the case of dative verbs, *to*-PPs would be VP-attached as recipients. This attachment satisfies an argument position and is a high-frequency option. In the PennTreebank, 33% of the PPs modifying our dative verbs were recipients in our judgment, and in Boland and Boehm-Jernigan's (1998) sentence completion data from fragments such as *Paul gave the script* . . . , a recipient argument was elicited 43% of the time. In the case of our action verbs, *with*-PPs would also be VP-attached as instruments, perhaps because the structure is lexically specified, but also because our normative data suggest that it is the most likely interpretation for readers to assign at that point.

In spite of producing the same outcome as syntax-first models, our materials do test the generality of a specific finding from the constraint-based literature concerning when plausibility effects should occur. This finding, made by Garnsey *et al*. (1997), is that when verb bias is strong, plausibility has no effect.

[6] Clifton *et al*. (1991) explicitly argue in their syntax-first reanalysis account that VP-adjunct attachment is attempted prior to any type of NP-attachment.

In their study, they used verbs that can take either a direct object or a sentence complement, and they grouped the verbs according to sentence complement frequency (estimated using sentence completion data). When they presented the verbs in sentence complement target sentences that contained temporarily ambiguous NPs, which were either plausible or implausible direct objects of the verb, direct object-biased verbs were difficult and sentence complement-biased verbs were easy regardless of plausibility. In contrast, equi-biased verbs were difficult only when the ambiguous NP was plausible as a direct object.

In our own sentence completion data, fragments from the action verb stimuli (truncated at the preposition *with*) were completed with unambiguous instruments 66% of the time. Fragments from the dative verb stimuli (truncated at the preposition *to*) were completed with animate, recipient NPs 75% of the time. While that difference may not seem large, sentence completions are not the only measure of verb bias. For example, there was a sharp contrast in our corpus data between the proportion of times that PPs modifying our action verbs were assigned the instrument role (5%) and the proportion of times that PPs modifying our dative verbs were assigned the recipient role (33%). Thus, under a verb-bias account, both these measures predict that the detection of an implausible recipient NP would be delayed relative to an implausible instrument NP.

Lastly, delay models, such as the restricted interactive model (Britt, 1994), predict no difference in the timing of plausibility detection for the PPs here. According to this model, PP-attachment decisions are delayed until the entire constituent has been constructed and the discourse model consulted if the verb does not *require* an argument. Indeed, all of our instrument verbs and all but one of our dative verbs (*handed*) do not require an argument. Thus, the attachment process and subsequent implausibility detection should be the same in all cases. Whether this process occurs quickly enough for the effect to appear at the NP is unclear, although Britt's own PP-attachment effects, which appeared after the embedded NP, suggest otherwise.

Table I summarizes the predictions regarding the timing of implausibility detection when the preposition is consistent with assignment of a recipient or instrument thematic role. To investigate these hypotheses, we investigated sentences like

Table I. Predicted Locus of Implausibility Detection Relative to the Implausible Embedded NP

	Dative PPs	Instrument PPs
Syntax-first model	Immediate	Immediate/absent/delayed
Verb-bias account	Delayed	Immediate
Delay model	Delayed	Delayed

NP, noun phrase; PP, prepositional phrase.

those in (1) and (2) using two different self-paced, word-by-word reading paradigms, reported in Experiments 1 and 2, below. We chose self-paced reading because it ensures that participants read and respond to the preposition; they cannot preview the preposition parafoveally, as readers might do in an eyetracking paradigm, and they are less likely to skip over it. While one might argue that eyetracking reflects more natural reading processes, it requires region-by-region analyses that may not capture the incremental processing occurring at the preposition or embedded NP.

OFFLINE NORMS

Materials

The stimuli included 20 dative PP items (10 plausible, 10 implausible) taken from Boland and Boehm-Jernigan (1998; Experiments 1 and 2) and 36 instrument PP items. All critical items are presented in the Appendix. We manipulated the consistency of the preposition and the plausibility of the noun in the first PP following the verb (PP1) in a two-by-two design. In all 36 instrument items, PP1 was headed by a preposition that was either consistent (*with*) or inconsistent (*of*) with an instrument interpretation. The object of the preposition was either a plausible (*glue*) or an implausible (*knobs*) instrument of the verb. A second PP headed by *with* (PP2) followed PP1 and always contained the real instrument phrase.

The content of the critical PP in the dative items was manipulated in the same way as the instrument items. The preposition in PP1 was either consistent (*to*) or inconsistent (e.g., *of, about*) with the recipient argument of the verb. In 10 items, the object of the preposition was a good recipient, and in 10 items it was a poor recipient. We randomized the 20 dative and 36 instrument items with 104 filler sentences to create four lists. Each item appeared only once per list, and conditions were balanced within and across lists. The dative items were matched, such that if participants saw the *to*-PP with the consistent recipient (1a), they saw the *of*-PP with the inconsistent recipient (1d), and vice versa. There were nine items per cell for the instrument PPs and five items per cell for the dative.

Because the dative stimuli were extensively normed in Boland and Boehm-Jernigan (1998), we collected normative data almost exclusively for the instrument materials. The exception is the set of dative sentence completions referred to in the introduction and described below.

Length and Frequency

To ensure that the target sentences were evenly matched across the four instrument conditions, we calculated the frequencies (Francis & Kucera, 1982)

and mean lengths of the only words that differed across conditions: the preposition and noun in PP1. The plausible instruments (e.g., *glue*) had a mean frequency of 36.39 and a mean length of 6.25 characters. The implausible instruments (e.g., *knobs*) had a mean frequency of 62.47 and a mean length of 6.08. The plausible and implausible nouns did not differ on length or frequency [$t(70) = 0.44$, $p > 0.10$; $t(70) = 1.43$, $p > 0.10$]. *With* is two characters longer than *of*, and it is about five times less frequent (*with* 7289 vs. *of* 36,411).

Instrument Plausibility

Eighteen Rutgers University undergraduates rated truncated but complete sentences, such as *The repairman sealed a container <u>with clear glue</u>*, on a scale from 1 (easy to perform the action indicated by the verb with the underlined noun phrase) to 7 (hard to perform . . .). There were two lists so that each participant saw each item in only one condition. There were no fillers. As expected, participants rated the plausible NPs (M = 2.45, SD = 0.93) as better instruments than the implausible NPs (M = 4.87, SD = 1.13) [$t(35) = 9.96$, $p < 0.01$].

Complex NP Ratings

We measured the relative acceptability of the complex NPs formed by the direct object of the verb and PP1 in a set of 42 items. These items used the four conditions shown in (2) and two additional conditions that substituted gerunds for the preposition and adjective (e.g., *a container holding glue/knobs*). We extracted the complex NP formed by the direct object and PP1/gerund out of each one and randomized them with 58 fillers that were composed of various complex NPs (NPs modified by PPs, gerunds, and *wh-* and *that* relative clauses) and a few NPs modified by adjectives. Each item appeared only once on each of six lists, and conditions were balanced within and across lists. Sixty Rutgers undergraduates judged how <u>natural</u> each complex NP was as a description of an object, event, or situation. Another 60 students judged how <u>likely</u> each was as an object, event, or situation. The scale always ranged from 1 (very unlikely or very unnatural) to 5 (very likely or very natural). Each participant completed only one of the 12 lists in a paper and pencil task that took approximately 20 minutes.

Table II gives the mean acceptability ratings for the 36 items that were used in Experiments 1 and 2. Six items were eliminated from the original set on the basis of the acceptability ratings and our own intuitions. In terms of naturalness, the mean scores for the four types of complex NPs were almost identical (*t*s < 1.1). However, the consistent plausible condition (*a container with clear glue*) was rated marginally more likely than the inconsistent implausible condition (*a container of clear knobs*) [$t(35) = 1.94$, $p = 0.06$].

Table II. Mean Naturalness and Likelihood Acceptability Ratings (with Standard Deviations) for the Subset of Critical Stimuli Used in Experiments 1 and 2[a]

Complex NP (n = 36)	Mean naturalness score (SD)	Mean likelihood score (SD)
a container with clear glue	3.13 (0.64)	3.38 (0.64)
a container of clear glue	3.00 (0.65)	3.23 (0.70)
a container with clear knobs	3.00 (0.76)	3.21 (0.68)
a container of clear knobs	3.06 (0.67)	3.07 (0.61)

NP, noun phrase; SD, standard deviation.
[a] 5 = very likely, very natural; 1 = very unlikely, very unnatural.

No other pairwise comparisons differed (ts < 1.6). Thus, the NP-attached interpretations of the critical PP were all equally acceptable and almost equally likely.

Noun Completions

To determine how frequently the direct object nouns in the critical materials select prepositions, we randomized each of the 36 direct objects to a list, along with 34 filler fragments of varying lengths, as the subject of an otherwise unfinished sentence (e.g., *A vase* ...). We collected completions from 15 Rutgers undergraduates. The critical subject NPs were all indefinites, just as in the online experiments. We tallied the proportion of PP completions (0.41) for the critical items and sorted them by preposition. Participants completed just over a third (0.35) of the critical fragments with a PP headed by *of*, but less than 1% (0.004) with a PP headed by *with*. The second most frequent preposition was *in* (0.04).

Instrument PP Completions

To measure how often the preposition *with* in PP1 is used as an instrument, we asked 18 Rutgers students to provide sentence completions for two types of sentence fragments created from the critical materials. The first type of fragment ended at the preposition (*The repairman sealed a container with* ...); the second ended at the immediately following adjective (*The repairman sealed a container with clear* ...). The 36 critical items were randomized with 64 filler fragments of varying lengths, and two lists were created. Each item appeared only once per list, and conditions were balanced within and across lists. We classified the responses into three categories: unambiguous instrument (61% of completions), unambiguous noninstrument (30%), and ambiguous (9%). Although the fragments ending at *with* seemed to elicit more instruments (66%) than the fragments ending in adjectives (56%), the difference between the arcsine transforms of the proportion data

was not statistically significant ($t < 1.1$). Thus, our materials were clearly biased toward an instrument interpretation for *with* in the critical PP.

Dative PP Completions

To measure how often the preposition *to* in PP1 is used as a recipient, we asked 36 Ohio State University students to provide sentence completions for truncated versions of the 20 critical sentences that ended at the preposition (*John gave a letter to . . .* , *Paul gave the script for . . .*). These 20 fragments were randomized with 50 filler fragments of varying lengths and presented as one of two lists. Each item appeared only once per list, and conditions were balanced within and across lists. Just as in the online experiment, those who saw the *to*-PP from the consistent recipient condition saw the *of*-PP from the inconsistent recipient condition and vice versa.

Participants completed 75% of the fragments that ended in *to* with an animate, recipient NP. For 14% of the fragments that ended in a different preposition, participants added a recipient PP headed by *to*. Thus, our materials were clearly biased toward a recipient interpretation for *to* in the critical PP.

EXPERIMENT 1

In Experiment 1, we investigated the timing of implausibility detection for recipient and instrument PPs and tested the competing predictions of constraint-based lexicalist models, syntax-first models, and delay models. Corpus frequencies and our sentence completion norms both suggested that our dative verbs were highly biased to take recipient *to*-PPs whereas our action verbs were less biased to take instrument *with*-PPs. Only the constraint-based, verb-bias account (Garnsey *et al*., 1997) predicted that implausible instrument NPs would be detected sooner than implausible recipients because of the influence of verb-bias. Syntax-first models (e.g., Frazier & Clifton, 1996) and delay models (e.g., Britt, 1994) either predicted no difference in timing or an advantage for implausible recipients.

Methods

Participants

Thirty-two Rutgers students participated for credit in an introductory psychology course. All were native English speakers.

Materials

This experiment used the materials described above in the section "Offline Norms."

Procedure

Participants were told that they would be reading individual sentences on a computer screen and that they should terminate any sentence as soon as it stopped making sense. In this "stop-making-sense" task, participants controlled the presentation rate of each word in a sentence by pressing a button labeled "go." Each press of "go" displayed the next word in the sentence in a cumulative fashion. Pressing a button labeled "stop" concluded a sentence and began the next trial. Sentences were presented on an IBM clone, and each one fit on a single line. All participants completed an eight-item practice session. The experiment took about 30 minutes.

Results

The primary data in this paradigm are the proportion of new "stop" responses at a given word position, and these were collected at 12 word positions beginning with the verb. We do not present the reaction times (RTs) to the "go" responses because pressing "stop" ended a trial before the end of the sentence and, as a result, led to incomplete data for later word positions. Mean latencies for "go" responses for the instrument data were about 440 ms at the fastest word position and about 900 ms at the slowest word position.

Figures 1 and 2 show the percentage of new "stop" responses (rather than the cumulative percentage) at each word position during the critical region for
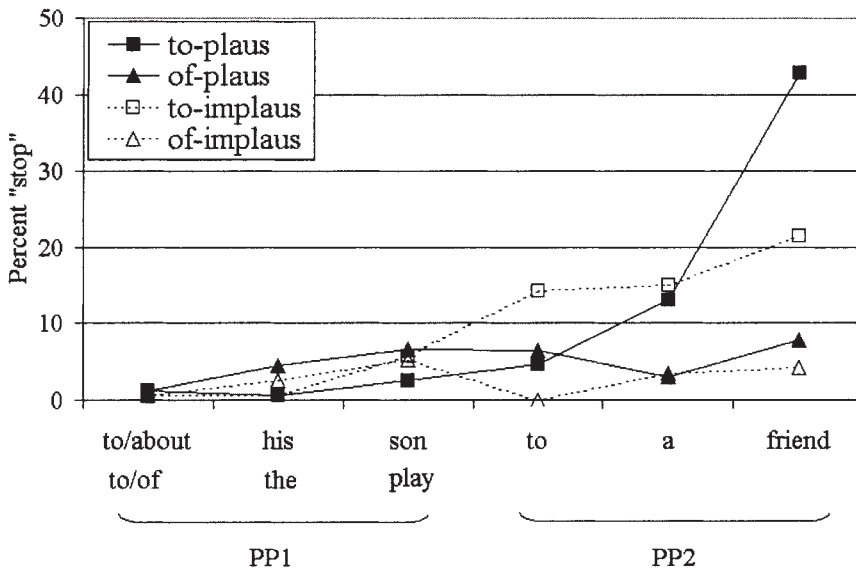

Fig. 1. The percentage of new "stop" responses in each condition, at each word position, for the dative stimuli of Experiment 1.
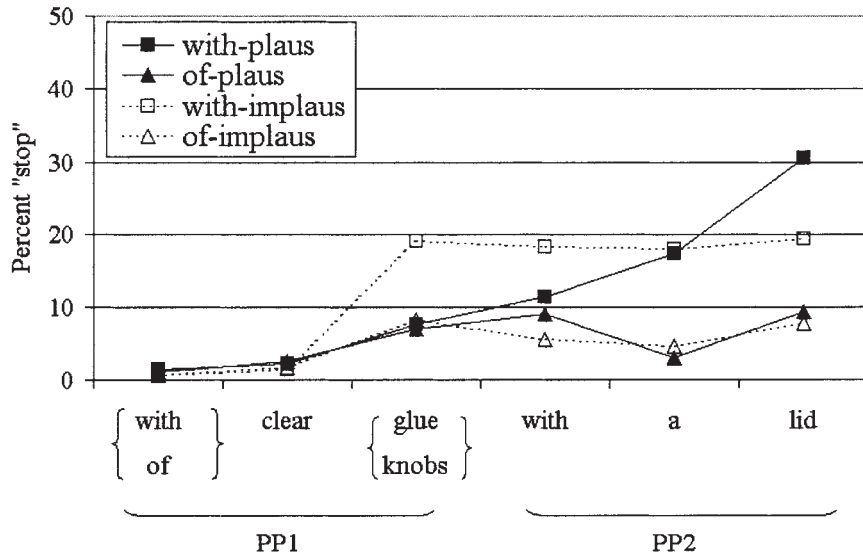
Fig. 2. The percentage of new "stop" responses in each condition, at each word position, for the instrument stimuli of Experiment 1.

the dative and instrument stimuli, respectively. We calculated these percentages in order to minimize the dependence between earlier values and later ones (as in Boland, Tanenhaus, & Garnsey, 1990, and Boland, Tanenhaus, Garnsey, & Carlson, 1995). We did not include the first three word positions because there were very few "stop" responses. The figures reveal some apparent differences between datives and instruments. First, although implausible recipient and implausible instrument nouns both cause a garden path effect when the preposition is consistent, they do so at different word positions. As in Boland and Boehm-Jernigan (1998), the effect of an implausible recipient was evident at the preposition in PP2. In the instrument conditions, however, an effect of plausibility emerged earlier, at the implausible NP itself. Second, there was a stronger garden path during PP2 in the consistent plausible condition for datives than for instruments.

To investigate these differences, we submitted the arcsine transforms of the percentages of new stop responses by participants to a 4(list) $\times$ 2(sentence type: dative or instrumental) $\times$ 2(plausibility: consistent or inconsistent with thematic role) $\times$ 2(preposition: consistent or inconsistent) $\times$ 6(word position: PP1 through PP2) repeated-measures analysis of variance (ANOVA). For the items analysis, we separated instruments from datives because plausibility was manipulated within items for the instrument stimuli and between items for the dative

stimuli. The instrument items were submitted to a 4(list) $\times$ 2(plausibility) $\times$ 2(preposition) $\times$ 6(word position) repeated-measures ANOVA. The dative items were submitted to a 2(list) $\times$ 2(plausibility) $\times$ 2(preposition) $\times$ 6(word position) repeated-measures ANOVA.[7]

We found a three-way interaction among plausibility, preposition, and word position [F1(5,140) = 5.50, $p < 0.01$], both for datives [F2(5,80) = 5.24, $p < 0.01$] and for instruments [F2(5,160) = 4.92, $p < 0.01$]. That is, for both sentence types, the garden paths in the consistent plausible condition and the consistent implausible conditions differed in time-course. Although there was no four-way interaction with sentence type (F1 = 1.02), sentence type interacted with preposition and word position [F1(5,140) = 3.30, $p < 0.05$]. To explore these interactions, we conducted additional ANOVAs at individual word positions beginning at the noun in PP1 because it was the first point at which an apparent effect was observed in Figure 2.

At the noun, there was a main effect of sentence type [F1(1,28) = 10.28, $p < 0.01$] that interacted with plausibility [F1(1,28) = 7.75, $p < 0.01$] and preposition [F1(1,28) = 6.75, $p < 0.05$]. In addition, plausibility and preposition interacted with each other [F1(1,28) = 7.33, $p < 0.05$], but only for the instrument PPs [F2(1,32) = 10.53, $p < 0.01$]. This suggests a garden path at the implausible noun (e.g., *knobs*) when the preposition was consistent with VP-attachment (*with*). Datives showed no reliable effects [Fs < 1.4].

At the preposition in PP2, there was a main effect of sentence type [F1(1,28) = 10.66, $p < 0.01$] with a higher percentage of "stop" responses in the instruments. There was also an interaction between plausibility and preposition [F1(1,28) = 16.57, $p < 0.01$] which was found both for datives [F2(1,16) = 16.52, $p < 0.01$] and for instruments [F2(1,32) = 8.79, $p < 0.01$]. Whereas the instruments had shown an effect of plausibility in the consistent preposition condition at the NP in PP1, a similar effect was not reliable in the datives until one word position later.

At the determiner, we found a main effect of preposition [F1(1,28) = 12.53, $p < 0.01$], both for datives [F2(1,16) = 13.77, $p < 0.01$] and for instruments [F2(1,32) = 32.45, $p < 0.01$]. Preposition did not interact with plausibility at this point (Fs < 1.0), suggesting that there was a garden path in all four consistent preposition conditions. At the noun, the effect of preposition interacted marginally with plausibility [F1(1,28) = 3.51, $p < 0.10$], for instruments [F2(1,32) = 2.80, $p < 0.10$] and datives [F2(1,16) = 3.4854, $p < 0.10$]. That is, conditions with a consistent preposition seemingly had a greater proportion of "stop" responses than the inconsistent conditions, presumably because PP1 was being

---

[7] Throughout, we report the Huynh-Feldt (Huynh & Feldt, 1976) adjusted probability values for analyses involving three or more levels of word position or region as a factor. The degrees of freedom are unadjusted.

reanalyzed, but the plausible consistent condition had the greatest proportion of all. Although there appeared to be a difference in magnitude between the dative and instrument conditions, there was no main effect of sentence type (F1 = 1.03) nor a three-way interaction between sentence type, plausibility, and preposition (F1 = 1.2).

Discussion

These data clearly show that the detection of an implausible recipient is delayed relative to the detection of an implausible instrument. Thus, these results are incompatible with syntax-first models, which predicted either that implausibility detection would occur at the embedded NP for both or that instruments would be delayed. These results are also inconsistent with delay models, which predicted 1. that implausibility detection would occur at the same time for both, and 2. that implausibility detection would occur after the embedded NP.

The results support the predictions of the verb-bias account. The data suggest that because dative *to*-PPs satisfy high-frequency argument structures, an implausible recipient is not detected until additional constraint information, the arrival of the second *to*, is processed. In contrast, because instrument PPs occur with our action verbs much less frequently than the recipient PPs occur with our dative verbs, verb-bias does not interfere with implausibility detection in the instrument conditions. While these data clearly rule out several parsing models, they do not indicate whether frequency interacts with argument status because instruments may or may not be arguments.

In spite of the differences between instrument and recipient PPs in the timing of implausibility detection, the dative verb data replicated the pattern and time-course of Boland and Boehm-Jernigan (1998; Experiments 1 and 2). In all three experiments, a late garden path in the implausible consistent condition demonstrated that PP1 was attached to the VP and assigned the thematic role of recipient when the preposition was consistent with the argument structure of the verb. If PP1 had not been attached as a dative argument at the preposition, the implausible noun should not have generated processing difficulty. Thus, we have added to the evidence that attachment decisions are made at the preposition for dative PPs and now for instrument PPs as well. Together, the Boland and Boehm-Jernigan data and the current dative data also demonstrate that appropriate NP-attachments were made for the *about* and *of* PPs as rapidly as VP-attachments with absolutely no signs of reanalysis, thus providing an additional piece of evidence against the claim that all PPs are initially VP-attached in accordance with minimal attachment (Frazier & Clifton, 1996; Rayner *et al.*, 1983).

In contrast to the consistent demonstration that the dative garden path begins at the preposition in PP2 and not at the noun in PP1 (the current experiment;

Boland & Boehm-Jernigan, Experiments 1 and 2), we have just one demonstration (the current experiment) of an earlier garden path effect for instrument items. Because the earliness of the effects suggests that instrument PPs were analyzed differently than dative PPs, it is important to determine whether or not the difference is meaningful. In Experiment 2, we will attempt to replicate the instrument effect and its time-course using a moving window, self-paced, reading time measure that, unlike stop-making-sense, does not have a decision component that elevates RTs. This is a very conservative test because moving window, self-paced reading is often noted to produce effects that lag a word or two behind the expected word position (e.g., Boland & Boehm-Jernigan, 1998; Taraban & McClelland, 1988). If we can replicate the early onset of the instrument garden path with the moving window task, it will provide strong evidence that the difference is, in fact, reliable and meaningful.

EXPERIMENT 2: INSTRUMENT PPS AND THE EARLY GP

The primary goal of this experiment was to test the reliability of the early garden path effect observed for the instrument sentences in Experiment 1. It is important to replicate the data pattern in an experimental paradigm that does not involve any explicit judgments and in which reading rates approach those in normal reading. We achieved this by using the moving window, self-paced reading paradigm (which, as we mentioned, is a conservative test because moving window often records effects one or two words downstream from the source) and by adding a contextual sentence to each sentence in the experiment. We expected that the contextual sentence would improve the overall naturalness of the sentences and facilitate reading rates. Thus, concern that presentation of these sentences out of context might be the true source of the garden paths in Experiment 1 should be mitigated.[8]

Methods

Materials

The stimuli consisted of the 36 critical instrument PP items from Experiment 1, plus a context sentence that was appropriate for all conditions in a given item. By appropriate, we mean that the sentence did not introduce any of the NPs or actions mentioned in the target sentence. For example, the contextual sentence for the example in (2) was *The renovation had taken more than*

---

[8] We would add that speakers do in fact produce two PPs in a row with the same preposition in natural speech. For example, a local Columbus, Ohio, newscaster asked, "Why would anyone want to kill this man with a kind heart with so many friends?"

a week to complete. Two additional conditions that substituted gerunds for the preposition and adjective (e.g., *a container holding glue*) also acted as fillers. The gerunds were intended to function as unambiguously NP-attached phrases, but their syntactic complexity resulted in long RTs that made comparison impossible. Although we excluded these conditions from all analyses, they have the unfortunate consequence of reducing power in our item analyses. The critical items were randomized with 80 pairs of filler sentences of varying syntactic complexity and were rotated across six lists. Each list contained 36 yes/no comprehension questions, none of which focused on PP1 or PP2.

Participants

Eighty-four students from the same pool as Experiment 1 completed the experiment.

Procedure

Participants were instructed that they would be reading pairs of sentences on a computer screen and that they would occasionally have to answer a yes/no question about the last pair they had read. Sentences were presented on an IBM clone one word at a time in a "moving window" design. Participants pressed a button labeled "start" to bring up a pair of sentences in which each letter in each word was covered by a dash-mark. They controlled the rate at which they read each word by repeatedly pressing a button marked "next." With each press of the "next" button, the preceding word (if there was one) returned to dash-marks, and the next word in the sentence was revealed. The experiment took about half an hour. Participants began by completing a practice session of 10 pairs of sentences complete with comprehension questions.

Results

Reading times (RTs) were collected for 12 word positions in the critical sentences, beginning at the verb. Response times greater than 2.5 standard deviations from a participant's mean were replaced with the cutoff value. Mean RTs from the onset of the PPs through the end of the sentence are shown in Figure 3. As expected, RTs in this paradigm were faster than "go" latencies in Experiment 2. Based on the results of Experiment 1, we conducted four individual ANOVAs, one at each critical word position starting with the noun in PP1.

At the noun (*glue/knobs*), there was the crucial interaction between plausibility and preposition [$F1(1,78) = 5.17$, $p < 0.05$; $F2(1,30) = 6.84$, $p = 0.01$]. Reading times for the *with*-implausible instrument condition were slower than the other three conditions, as in Experiment 1. Unlike Experiment 1, plausibility and preposition did not interact at any other word position ($Fs < 1$). The effect
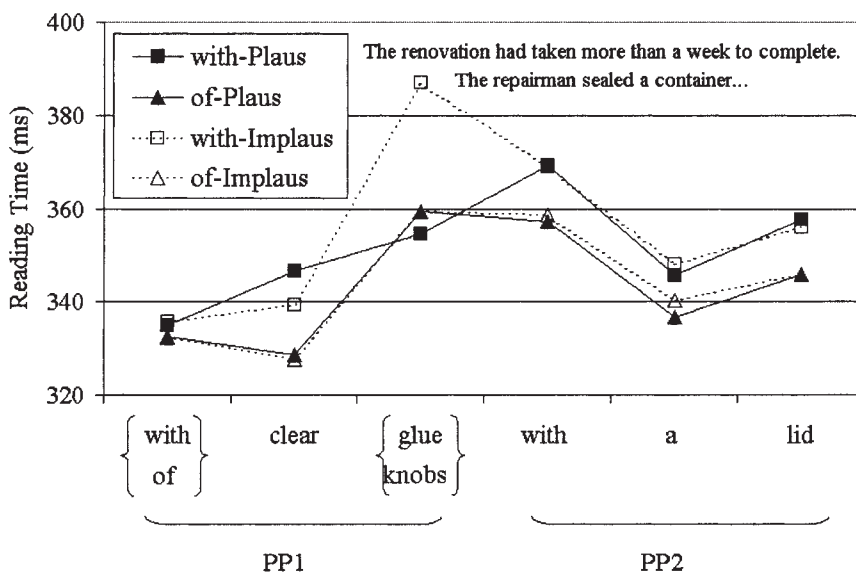
Fig. 3. Mean reaction times by word position for the four conditions in Experiment 2.

of preposition was marginal by items at the preposition [F1(1,78) = 2.618, $p = 0.11$; F2(1,30) = 4.065, $p = 0.05$] and by subjects at the determiner [F1(1,78) = 3.580, $p = 0.06$; F2(1,30) = 2.832, $p = 0.10$]. The effect was not fully reliable until the noun, and then only by subjects [F1(1,78) = 5.223, $p < 0.05$; F2(1,30) = 3.337, $p < 0.10$].

Discussion

Once again, we obtained evidence of the early instrument garden path effect in the consistent implausible condition; it occurred at the implausible noun, replicating the timing obtained in Experiment 1 and demonstrating that the instrument garden path is robust. Furthermore, this again demonstrates that PP-attachment decisions for instrument PPs were made at the preposition.

Although replicating the interaction between plausibility and preposition at the noun in PP2 was not our primary motivation for this experiment, we were surprised that it did not occur. One possible explanation is that the difficulty associated with processing two *with*-PPs in a row, even when the first one can be reassigned a different thematic role and/or attachment, masked the interaction. Alternatively, the pattern may demonstrate that the stop-making-sense task is more sensitive to small local garden paths than the moving window task. This contrast

between tasks was noted in the discussion of Boland and Boehm-Jernigan (1998). On the other hand, the moving window task appears to be more sensitive to frequency effects. High-frequency *of*-PPs were significantly faster than lower frequency *with*-PPs [$F1(1,78) = 12.027$, $p < 0.01$; $F2(1,30) = 7.902$, $p < 0.01$] at *clear* in Experiment 2. This contrast was not observed in Experiment 1.

GENERAL DISCUSSION

In Experiment 1, we found differences in the timing of the detection of implausible dative and instrument PPs. Implausible instruments were detected immediately at the NP whereas implausible recipients did not increase processing load right away. In fact, a garden path in the implausible recipient condition was not detected until a second cue to misanalysis arrived; namely, the onset of a second *to*-PP. The relatively late detection of implausible recipients replicated the pattern of effects seen in Experiments 1 and 2 of Boland and Boehm-Jernigan (1998), and the relatively early detection of implausible instruments was replicated in the current Experiment 2. On the basis of these results, we argued against syntax-first models, which predicted that there would be no difference in the timing of implausibility detection (or that there would be a delay for instruments), and we argued against delay models, which predicted no difference in timing or that both would be detected after the implausible NP.

Our results are most consistent with the verb-bias account of Garnsey *et al.* (1997) and with constraint-based models in general. This is because constraint-based models already contain a mechanism by which implausible NPs can escape detection. When verb bias is strong, plausibility will have a limited effect by itself and will require additional constraint information to support the needed semantic and syntactic alternative. More specifically, dative *to*-PPs satisfy high-frequency argument positions that many linguists and psycholinguists would agree are likely to be represented in the lexical entries of dative verbs. According to constraint-based models, a reader or listener would activate the lexical entry of the verb and the argument structure alternatives that are encoded with it. These syntactic alternatives would be weighted by frequency with recipient *to*-PPs representing strong competitors, and the arrival of the preposition would add further support to that alternative. Thus, a high level of commitment would be afforded to that syntactic structure and thematic role assignment beginning at the verb. It is not surprising then that the onset of an implausible recipient would not be enough to override a high-frequency lexical commitment that originated at the verb and to cause implausibility detection.

In the case of instrument *with*-PPs, several detailed explanations for the early garden path effect are possible depending on the constraint-based model of interest. For example, some models include lexical representations for all

attachments regardless of argument status (MacDonald *et al.*, 1994). In contrast, the argument structure hypothesis (ASH) of Lewis and Boland (2000) includes lexical representations for argument attachments and globally generated representations for adjunct attachments. If *with*-PPs are represented lexically, then the *with*-attachment site would be generated lexically at the verb and be available for an immediate attachment decision at the preposition; there would be no difference in the timing of attachment decisions for dative and instrument PPs. However, because instrument attachments would be generated from weaker lexical representations than dative attachments (i.e., instrument *with*-PPs occur with action verbs less frequently than *to*-PPs occur with dative verbs), these weaker representations would be more susceptible to alternative constraint information.

In contrast, if instrument PPs are adjuncts (despite their behavior on traditional linguistic tests of argument status) and thus not represented lexically, according to the argument structure hypothesis, then NP and VP adjunct structures would be generated globally at the preposition and multiple constraints would be used to guide ambiguity resolution. Based on our sentence completion data, there is a strong constraint to treat a *with*-PP following an action verb as an instrument. While this might appear to be similar to the dative stimuli in that there is a strong constraint in favor of a certain syntactic and thematic assignment, ASH allows us to make a distinction between lexically generated argument attachments and globally generated adjunct attachments, which might, in addition to the frequency effects, explain the difference in timing. It follows from ASH that even though the decision to attach a preposition into the existing structure would occur at the preposition for both datives and instruments, a globally generated PP-attachment site would be available at a slightly later time-point than a dative PP-attachment site that was made available during lexical access at the verb. For this reason, ASH predicts that instrument PPs would be less resistant to the influence of other constraint information than dative PPs. This explanation is similar, in part, to the Ferreira and Henderson (1995) finding that the further away a disambiguating word was from its head (i.e., the more time that had passed since the initial attachment), the harder it was to reanalyze the structure.

Although these data can be accounted for within more than one constraint-based lexicalist model, they are clearly inconsistent with syntax-first models. More specifically, they are inconsistent with claims of a thematic processor, which establishes the most likely thematic relations and then checks those relations against the initial structure (Clifton *et al.*, 1991; Frazier, 1990; Rayner *et al.*, 1983). The data suggest that plausibility relations are not always immediately available to trigger reanalysis and, somewhat counter-intuitively, that this is true of the high-frequency argument relations. One might expect that a thematic processor that immediately checks thematic relations against structure would find violations of argument roles to be the most robust indicators of an anomaly.

Explanations of reanalysis processes in syntax-first models have remained largely untested, perhaps because they are treated as secondary to demonstrations of minimal attachment and late closure effects. However, the current data demonstrate that the thematic processor and its anomaly detection and reanalysis processes require empirical investigation in their own right if we are to have a thorough account of human sentence processing behavior.

To be fair, these data also suggest that constraint-based models have yet to adequately model, predict, and test argument/adjunct differences, although these issues may be especially challenging given that it is difficult to find argument and adjunct phrases that do not also contrast in frequency. Nonetheless, we feel that argument status is an important area for continued constraint-based research because of its role in limiting the power of these approaches.

APPENDIX

Listed below are the critical items for Experiments 1 and 2. Items 1 to 36 contain the instrument stimuli. The first sentence in each is the contextual sentence for Experiment 2. Preposition and noun were crossed in a two-by-two design. The first noun in each pair is the plausible instrument. Items 37 to 57 contain the dative stimuli. In items 37 to 46, the target noun is consistent with a recipient role. In items 47 to 56, the target noun is inconsistent with a recipient role.

1. The lawn of the old Victorian needed a lot of work.
   The landscaper lifted a bag {with/of} large {shovels/shrubs} with a forklift to make room.
2. Getting everything into the barn was top priority.
   The hunter moved a sack {with/of} leather {straps/hides} with a tractor while snow fell.
3. The sale at Home Depot had been a success.
   The manager opened a carton {with/of} garden {shears/hoses} with a razor to help restock.
4. The greenhouse wasn't always a very relaxing place.
   The gardener smashed a pot {with/of} sandy {rocks/dirt} with a spade as fans whirred.
5. The old apartment wasn't the nicest place to move into.
   The woman broke a shelf {with/of} heavy {dishes/brackets} with a hammer while she unpacked.
6. The customs area was downright chaotic.
   The deputy tagged a chest {with/of} official {stamps/uniforms} with a laser as he worked.
7. The equipment shed always seemed to need reorganizing.

The surveyor marked a can {with/of} white {chalk/bolts} with a label while griping aloud.

8. The coffeeshop was noisy and crowded.
The busboy stacked a counter {with/of} dirty {cups/flies} with hot plates in the back.

9. The clients would be visiting the agency they had hired.
The marketer designed an ad {with/of} colored {pencils/squares} with a computer as a demo.

10. The renovation had taken more than a week to complete.
The repairman sealed a container {with/of} clear {glue/knobs} with a lid while he smiled.

11. The local market was bustling with shoppers.
The butcher wrapped a package {with/of} thick {napkins/steaks} with a cloth as he gabbed.

12. The CEO knew his office was in good hands whenever he was away.
The secretary closed a packet {with/of} extra {plastic/fliers} with a seal to be safe.

13. Going from store to store was a tough way to earn a living.
The vendor bound a parcel {with/of} tan {elastics/folders} with some twine for a handle.

14. The vacant lot was a fun place to play.
The boy wound a bundle {with/of} sharp {wire/sticks} with some rope while telling jokes.

15. Earnings for the upstart auction house were poor.
The firm taped an auction {with/of} old {equipment/furniture} with a camcorder to study why.

16. The trip wasn't going to be much fun in the rain.
The scout sheltered a crate {with/of} dry {canvas/wood} with a tarp to help out.

17. The building near headquarters was an easy target.
The spy saw a box {with/of} black {binoculars/grenades} with a telescope while he waited.

18. There was a bulletin board at the end of each aisle in the hardware store.
The clerk hung a photo {with/of} silver {tacks/keys} with some pushpins at aisle one.

19. Designing his own home had been a life long dream.
The builder supported a ceiling {with/of} new {beams/stucco} with a pole to add force.

20. The studio was a quiet and relaxing place to work.
The artist created a pattern {with/of} pastel {paints/trees} with a stencil as it rained.

21. The new industrial park was still under construction.

The engineer layered a bin {with/of} flexible {tubing/drains} with a screen to save time.
22. The institution was practically empty in the early morning.
The janitor cleaned a closet {with/of} harsh {solvents/odors} with a mop to pass time.
23. The condemned building was being demolished.
The foreman raised a vault {with/of} strong {levers/cement} with a pulley as trucks idled.
24. The construction site attracted mischief.
The vandal damaged a wall {with/of} red {bricks/doors} with a backhoe for no reason.
25. The patient had arrived at the clinic right on time.
The doctor emptied a jar {with/of} sterile {tweezers/specimens} with latex gloves at a desk.
26. Mother's Day was one of the busiest holidays of the year.
The florist arranged a vase {with/of} clear {marbles/water} with tea roses while she sat.
27. Calling a plumber would have been very costly.
The barber unclogged a sink {with/of} tough {chemicals/hairballs} with a plunger on his own.
28. The mansion would be closed for the winter months.
The maid protected a basket {with/of} pink {blankets/soaps} with a sheet while mice hid.
29. The shed in the back yard had been converted into a professional workshop.
The carpenter divided a section {with/of} deep {panels/scratches} with a ruler as saws ran.
30. The warehouse was a confusing place of bodies and shipments.
The worker transported a locker {with/of} sturdy {rollers/tiles} with a cart to the ramp.
31. Everyone was glad to be part of the campaign.
The designer drew a layout {with/of} blue {markers/dresses} with new software for the ad.
32. The lamp didn't seem to be working when needed.
The electrician secured a line {with/of} copper {wiring/switches} with some tape as a fix.
33. Taking inventory wasn't a very exciting task.
The agent labeled a case {with/of} orange {tags/toys} with his name while chewing gum.
34. Many of the science clubs were allowed to use the facilities after school.
The student preserved a sample {with/of} pure {vinegar/liquid} with a freezer in a lab.
35. The front porch was a warm and sunny place to work.

The retiree mended a vest {with/of} metallic {thread/fabric} with a needle as she relaxed.

36. The emergency room had more patients than they could handle.
The nurse cleared a tray {with/of} bloody {forceps/bandages} with clean tongs to tidy up.

37. A lawyer revealed the payment {to/of} the witness to a client during the trial.

38. The man sent the response {to/of} his boss to the office by fax machine.

39. The coach mailed a message {to/about} the team to the school after the game.

40. John gave a letter {to/about} his son to a friend a month ago.

41. The farmer showed the gate {to/of} the village to a boy during the festival.

42. Cher promised a wave {to/toward} the crowd to her manager at each appearance.

43. The principal delivered a reminder {to/about} the kids to their moms at a meeting.

44. Anna handed a note {to/about} the teacher to the child right after lunch.

45. Ruth taught the instructions {to/about} the staff to the manager at the store.

46. The assistant offered the advice {to/about} the choir to the pastor a week ago.

47. The dealer revealed the rules {to/of} the game to the player at the club.

48. The poet sent the words {to/of} the poem to a critic after the lecture.

49. The author mailed a conclusion {to/of} the book to the editor after a delay.

50. Paul gave the script {to/of} the play to a girl during the break.

51. A guard showed the map {to/of} the museum to a lady after closing time.

52. Howard promised a resolution {to/of} the conflict to the diplomat on the sly.

53. The trucker delivered the lids {to/of} the jars to the factory in the morning.

54. Jack handed the key {to/in} the cage to the chimp during the experiment.

55. The tutor taught the tune {to/of} the song to his pupil before the test.

56. An engineer offered one solution {to/of} the problem to the company in an hour.

REFERENCES

Boland, J. E. (1997). Resolving syntactic category ambiguities in discourse context: Probabilistic and discourse constraints. *Journal of Memory and Language*, *36*, 588–615.

Boland, J. E., & Boehm-Jernigan, H. (1998). Lexical constraints and prepositional phrase attachment. *Journal of Memory and Language*, *39*, 684–719.

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

Boland, J. E., Tanenhaus, M. K., Garnsey, S. M., & Carlson, G. N. (1995). Verb argument structure in parsing and interpretation: Evidence from *wh*-questions. *Journal of Memory and Language*, *33*, 251–283.

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

Britt, M. A., Gabrys, G., & Perfetti, C. A. (1993). *A restrictive interactive model of parsing*. In Conference Proceedings of the 15th Annual Conference of the Cognitive Science Society (pp. 260–265). Hillsdale, NJ: Erlbaum.

Cinque, G. (1999). *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. New York: Oxford University Press.

Clifton, C., Jr., Speer, S. R., & Abney, S. P. (1991). Parsing arguments: Phrase structure and argument structure as determinants of initial parsing decisions. *Journal of Memory and Language*, *30*, 251–271.

Ferreira, F. N., & Henderson, J. M. (1995). Reading processes during syntactic analysis and reanalysis. In J. M. Henderson, M. Singer, & F. Ferreira (Eds.), *Reading and Language Processing* (pp. 119–147). Mahwah, NJ: Lawrence Erlbaum.

Francis, W. N., & Kucera, H. (1982). *Frequency Analysis of English Usage: Lexicon and Grammar*. Boston: Houghton Mifflin.

Frazier, L. (1978). *On Comprehending Sentences: Syntactic Parsing Strategies*. Ph.D. dissertation, University of Connectieut.

Frazier, L. (1990). Parsing modifiers: Special purpose routines in the human sentence processing mechanism? In D. A. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds.), *Comprehension Processes in Reading* (pp. 303–330). Hillsdale, NJ: Erlbaum.

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

Garnsey, S. M., Pearlmutter, N. J., Myers, E., & Lotocky, M. A. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language*, *37*, 58–93.

Huynh, H., & Feldt, L. S. (1976). Estimation of the box correction for degrees of freedom from sample data in randomized block and split-plot designs. *Journal of Educational Statistics*, *1*, 69–82.

Lewis, R. L., & Boland, J. E. (2000). *In Search of Fully Lexical Parsing*. Unpublished manuscript, University of Michigan.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). Lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*(4), 676–703.

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

Schutze, C. T. (1995). PP attachment and argumenthood. *MIT Working Papers in Linguistics*, *26*, 95–151.

Schutze, C., & Gibson, E. (1999). Argumenthood and English prepositional phrase attachment. *Journal of Memory and Language*, *40*, 409–431.

Sedivy, J., & Spivey-Knowlton, M. (1994). The use of structural, lexical, and pragmatic information in parsing attachment ambiguities. In C. Clifton, Jr., L. Frazier, & K. Rayner (Eds.), *Perspectives on Sentence Processing* (pp. 389–413). Hillsdale, NJ: Lawrence Erlbaum.

Speer, S. R., & Clifton, C., Jr. (1998). Plausibility and argument structure in sentence comprehension. *Memory and Cognition*, *26*(5), 965–978.

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

Taraban, R., & McClelland, J. L. (1988). Constituent attachment and thematic role assignment in sentence processing: Influences of content-based expectations. *Journal of Memory and Language*, *27*, 597–632.

Traxler, M. J ., Pickering, M. J., & Clifton, Jr., C. (1998). Adjunct attachment is not a form of lexical ambiguity resolution. *Journal of Memory and Language*, *39*, 558–592.

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