

RUNNING HEAD: AVOIDING ATTACHMENT AMBIGUITIES

Avoiding Attachment Ambiguities: the Role of Constituent Ordering

Jennifer E. Arnold

University of Rochester

Thomas Wasow, Ash Asudeh,

Stanford University

Pete Alrenga

University of California at Santa Cruz

Address correspondence to:

Jennifer Arnold

University of Rochester

Dept. of Brain and Cognitive Science

Meliora Hall 495

RC Box 270268

Rochester, NY 14627

ABSTRACT

Three experiments investigated whether speakers use constituent ordering as a mechanism for avoiding ambiguities. In utterances like "Jane showed the letter to Mary to her mother", alternate orders would avoid the temporary PP-attachment ambiguity ("Jane showed her mother the letter to Mary", or "Jane showed to her mother the letter to Mary"). A preference judgment experiment confirmed that comprehenders prefer the latter orders for dative utterances when the former order would have contained an ambiguity. Nevertheless, speakers in two on-line production experiments showed no evidence of an ambiguity avoidance strategy. In fact, they were slightly more likely to use the former order when it was ambiguous than when it was not.. Speakers' failure to disambiguate with ordering cannot be explained by the use of other ambiguity mechanisms, like prosody. A prosodic analysis of the responses in Experiment 3 showed that while speakers generally produced prosodic patterns that were consistent with the syntactic structure, these patterns would not strongly disambiguate the PP-attachment ambiguity. We suggest that speakers do not consistently disambiguate local PP-attachment ambiguities of this type, and in particular do not use constituent ordering for this purpose. Instead, constituent ordering is driven by factors like syntactic weight and lexical bias, which may be internal to the production system.

To what extent do speakers design the form of their utterances for their addressees? Some choices are clearly dependent on the need for clear communication, for example the choice of language (in a bilingual community, for instance), or the decision to speak loudly if the addressee is far away. But many production decisions may proceed independently of a consideration of either specific or general properties of the listener. A full understanding of the production system will need to explain which processes, if any, are addressee-oriented.

A specific instance of this question is whether speakers avoid ambiguous sentence structures. For example, a speaker who plans to say (1) may realize that it contains a local PP-attachment ambiguity, or at least that it may be difficult to understand.

(1) The judge sent the letter to the president to the members of the congressional subcommittee.

Of course all sentences contain ambiguity at numerous levels, for example the beginning of the word “president” is consistent with cohort competitors like “present,” and “The judge sent the letter” could be a reduced relative clause. However, here we only focus on local PP-attachment ambiguities, which have been shown to sometimes cause parsing difficulty (e.g., Altmann & Steedman, 1988; Boland & Bohem-Jernigan, 1998; Frazier, 1978; Rayner, Carlson, & Frazier, 1983; Spivey-Knowlton & Sedivy, 1995), with potentially long-term consequences for comprehension (Christianson, Hollingworth, Halliwell, & F. Ferreira, 2001).

If speakers believe local ambiguities like this will cause difficulty for the listener, they may take pains to avoid them. Several potential mechanisms for this process have been proposed. The speaker may rephrase the utterance, for example “There’s a letter to the president, and the judge got it and sent it...”. In spoken language, the prosodic realization of the utterance has also been shown to facilitate the interpretation of PP-attachment ambiguities, whether the

speaker intended the prosody to disambiguate (Snedeker & Trueswell, 2003; Snedeker et al., 2000), or not (Warren, Shafer, Speer, & White, 2000).

A third, very natural mechanism for avoiding ambiguity is word order, the issue of interest here. For example, if the speaker wanted to convey the message evoked by the sentence in (1), the PP-attachment ambiguity could be avoided with the double-object construction: “The judge sent the members of the congressional subcommittee the letter to the president.” Indeed, writing manuals often cite word order as a desirable technique to avoid ambiguity (e.g. the National Archives and Records Administration’s guide to legal writing, 2003), and some theories have posited that ambiguity avoidance constrains the shape of grammars (Frazier, 1985; Hankamer, 1973). If speakers make ordering choices to avoid ambiguity, it would additionally be specifically relevant to a debate in the literature about whether choices in ordering are made to facilitate comprehension (Hawkins, 1994), or because of constraints on planning and production (Arnold, Wasow, Losongco, & Ginstrom, 2000; Stallings, MacDonald, & O’Seaghdha, 1998; Wasow, 1997).

We investigated this issue by asking whether speakers use phrasal ordering to avoid temporary PP-attachment ambiguities. Certain constructions in English allow variation in the order of the postverbal constituents, for example the Dative Alternation (see 2), where one order contains a local ambiguity, but the other orders avoid that ambiguity. We characterize these orders in terms of whether the goal argument comes relatively early in the utterance, since all constructions with the goal before the theme avoid the ambiguity.

(2) Examples of the Dative Alternation:

- a. theme-early (prepositional): Give the letter to Kim to me.
- b. goal-early (double-object): Give me the letter to Kim.
- c. goal-early (prepositional shifted): Give to me the letter to Kim.

The choice between orders in dative and other constructions has been shown to be influenced by a variety of factors. The most extensively documented determinant of postverbal ordering is the syntactic weight or complexity of the two constituents, such that shorter, “lighter” constituents tend to precede longer, “heavier” ones (see among others, Arnold, et al., 2000; Behagel, 1909/10; Hawkins, 1994; Stallings, et al., 1998; Wasow, 1997; but see Hawkins, 1994, and Yamashita and Chang, 2001, for evidence of the opposite pattern in Japanese). Ordering has also been shown to be influenced by the accessibility of the concepts referred to by each constituent, such that given or accessible arguments tend to precede new or inaccessible ones (Arnold et al., 2000; Bock & Irwin, 1980). Biases associated with particular lexical items have also been shown to play a role (Stallings et al., 1998; Wasow, 1997). However, none of these factors, together or independently, completely determines the order a speaker will choose. Therefore constituent ordering offers a possible mechanism for avoiding PP-attachment ambiguities, if speakers choose to do so.

The current study complements a number of recent studies that have debated whether ambiguity avoidance drives the use of other features of language production. For instance, two unrelated studies have shown that untrained speakers spontaneously produce instructions with prosody that disambiguates global PP-attachment ambiguities (Snedeker & Trueswell, 2003; Warren et al., 2000). Snedeker and Trueswell argue that such prosody is produced for the

purpose of disambiguation, and show that disambiguating prosody is only produced when the referential context does not disambiguate the meaning. By contrast, Warren et al. argue that instead it is a natural consequence of producing particular syntactic structures, and does not provide evidence of a disambiguation strategy. Similar findings have also been presented for local PP-attachment ambiguities (Kraljic & Brennan, 2003; Warren, 1985; Straub, 1997). Given these findings, in Experiment 3 we investigated whether speakers used prosody to disambiguate, using utterances like (2b) with two prepositional phrases. However, our study differs from all of the above studies, with the exception of Kraljic and Brennan (2003), in that they studied utterances that contained only a single prepositional phrase, for example “Tickle the frog with the flower” (Snedeker & Trueswell, 2003). By contrast, the presence of two prepositional phrases, as we describe below, reduces the availability of prosodic cues to NP-attachment of the ambiguous PP.

Ambiguity avoidance has also been investigated in a study of the production of optional words in sentences like *The chiropractor observed (that) {you / I} couldn't stand up straight* (V. Ferreira & Dell, 2000). In the absence of an explicit *that*, the embedded subject *you* initially could have been interpreted as either the direct object of “observe”, or the subject of a sentential complement. By comparison, “I” is morphologically constrained to be the subject of the following clause. In both cases the local ambiguity was disambiguated by the following context. In a series of experiments, V. Ferreira and Dell found that *that*-insertion was equal across these two conditions, leading them to conclude that ambiguity avoidance does not affect this choice. Instead, speakers were more likely to produce a *that* when the following information was less available, suggesting that optional word choice was driven by the needs of the production system.

A host of other studies have investigated the extent to which speakers take the addressee's perspective into account when designing referential expressions. On one hand, it has been argued that reference is established through a collaborative process between speaker and addressee, where the speaker takes the addressee's attention and state of knowledge into account dynamically (Clark & Wilkes-Gibbs, 1986; Clark & Krych, 2002). Even children as young as 5 have been shown to design referring expressions with respect to "common ground", or the information shared between them and their addressee (Nadig & Sedivy, 2000). On the other hand, some researchers have argued that common ground is only used at a second, monitoring stage of production, and not during initial utterance formulation (Horton & Keysar, 1996).

As the literature on this topic has shown, at issue is not simply whether speakers pay attention to the listener's needs (see Schober & Brennan, in press, for a review). Rather, a complete model of the production system will need to specify the circumstances under which speakers do or do not adapt to the needs of the addressee, and the precise processes and mechanisms at work. While the current study will not be able to answer the higher-order question, it contributes to the answer by addressing a plausible mechanism for avoiding local ambiguities – constituent ordering. The role (or lack of a role) that constituent ordering plays is suggestive of both how speakers order constituents and choose syntactic constructions, and the extent to which speakers consider ambiguity when planning an utterance. We will additionally provide evidence about the use of prosody in sentences like "The judge sent the letter to the president to the members of the congressional subcommittee."

The logic used in these experiments is to investigate the production of dative sentences that contain an ambiguity in the theme-early (prepositional) order, as in "Give the letter to Kim to me." These are compared with control sentences that contain no such ambiguity or where the

structural ambiguity is unlikely to cause a garden-path, like "Give the letter regarding Kim to me" and "Give the letter about Kim to me." If speakers tend to choose a goal-early order more often when the theme-early order contains an ambiguity than when it does not, it will be interpreted as evidence for an ambiguity avoidance strategy. In all experiments below, the relative size of the two constituents was held constant, so that weight was not a major determinant of constituent ordering (cf. Hawkins, 1994; Wasow, 1997).

In one preference judgment experiment and two on-line production experiments, we investigated ordering preferences in dative utterances like those in (2). All three experiments also investigated Heavy-NP-shift constructions, for example "The chef put the jello in the fancy mold in the refrigerator" or "The chef put in the refrigerator the jello in the fancy mold," which also contain an ambiguity in the unshifted order that can be avoided with the shifted order. However, the shifted version of this construction is highly dispreferred, and none of our experimental manipulations did much to modulate this dispreference in either the preference judgment task or the spoken production task. Therefore these data will be excluded from further discussion.

Our first experiment investigated whether comprehenders prefer goal-early orders when they avoid an ambiguity. In a forced-choice questionnaire, participants were asked to read pairs of sentences such as those in Table 1 (below), and choose the one they preferred. We manipulated "potential ambiguity", that is, whether an utterance would be ambiguous if it occurred in the theme-early order. Experiments 2 and 3 investigated ordering preferences in the on-line production of the same kind of utterance.

EXPERIMENT 1

Method and Materials

Participants were presented with 8 Dative experimental stimuli, randomly ordered with 16 fillers and 8 Heavy-NP-shift stimuli in 8 presentation lists. In all experimental items the two constituents contained roughly the same number of words. The full set of experimental dative stimuli are shown in Appendix A. Ambiguity was counterbalanced across lists and participants, such that half of the items that each participant saw contained a potential ambiguity, and each item occurred equally in the two ambiguity conditions across lists. The first three items were fillers, and were the same for every list. The two sentences in each pair were identical except for the relative order of the two constituents. Half the items for every questionnaire had the goal-early (double object) version first, half had the theme-early (prepositional) version first. Each item appeared with the goal-early order first on half the lists, and with the theme-early order first on the other half.

[INSERT TABLE 1 ABOUT HERE]

Participants were asked to read each pair carefully, and decide which one sounded better. They were asked to take the task seriously, but not to spend too much time on any one item. Most people finished the questionnaire in less than 20 minutes.

Participants

32 native English speakers from the Stanford community participated in exchange for \$7. Native English speaker was defined as having begun to learn the language before age 5.

Results and Discussion

The results show that as expected, readers chose the goal-early order more when it avoided an ambiguity that would have been present in the theme-early order (53%, SE = .07), compared with when it did not avoid any ambiguity (22%, SE = .05). This pattern is reflected in the results of analyses of variance, performed over participant and item means, which showed a main effect of ambiguity ($F(1,31) = 21.33, p < .001$; $F(1,7) = 24.98, p = 0.002$).

This finding is consistent with results from an on-line reading study reported by Boland and Boehm-Jernigan (1998). In the conditions most comparable to the present experiment, participants read sentences like “John gave a letter to/for his son to a friend a month ago.” Results showed that reading difficulty increased when the first PP contained a “to” compared to when it contained a “for”. Thus, on-line reading comprehension is impacted by this kind of temporary ambiguity, and replacing the preposition significantly improves the readability of the sentence.

It is evident, therefore, that comprehenders disprefer utterances with temporary PP-attachment ambiguities. Given the choice, they prefer constituent orders that avoid this ambiguity. The following experiment investigates whether speakers use constituent ordering to avoid ambiguous utterances of this type in on-line spoken sentence production.

EXPERIMENT 2

Methods and Materials

This experiment was designed to elicit utterances like those used in Experiment 1, but without biasing the speaker toward one construction or the other. Pairs of participants were

asked to participate together, one as speaker and one as addressee. Addressees were included to make the task more like a real conversation and lend ecological validity to the task.

The speaker was seated in front of a computer screen, and the addressee was given a pencil and a sheet of paper with questions. Each trial began when the speaker pressed a button, and a stimulus sentence like those in Table 2 appeared on the computer screen. In order to not bias the speaker toward a particular construction, the stimulus sentence presented the constituents of the target utterance in an order that avoided both the local ambiguity and the constructions we were trying to elicit, for example “A museum received Grant’s letters to Lincoln from the foundation.” Most sentences achieved this by passivizing the target utterance; 4 used verbs with an alternate argument structure (e.g., “received” instead of “sent”). The speaker was asked to read the sentence and remember its meaning as accurately as possible. When ready, the speaker pressed a button and the sentence disappeared. The computer also beeped, which prompted the addressee to read a sentence off the list, for example “What did the foundation do?” The speaker was instructed to not simply repeat the sentence they had read, but rather to answer the question, for example say “The foundation/ it gave....” We used the practice items to make sure they followed this practice, which elicited either the double-object or prepositional dative construction. The listener used a pencil to choose the correct of two alternative answers on the question/answer sheet.

The experimental items manipulated the potential ambiguity of the response – that is, whether the target utterance would be ambiguous in the theme-early construction or not. Sixteen dative items were presented in one of the two ambiguity conditions, such that each participant saw equal numbers of ambiguous and unambiguous items, and each item occurred equally in each condition across lists. The experimental items were pseudo-randomly combined with

sixteen heavy-NP-shift items (not discussed here) and 48 fillers. These were presented in 2 lists, each of which was presented both forwards and backwards. The same 10 practice items preceded each list.

[INSERT TABLE 2 ABOUT HERE]

The target utterances (i.e., what participants would produce when answering the question) were modified versions of the stimuli in Experiment 1, plus eight additional items of similar design (see Appendix A for the full set of stimulus sentences). Across all items the difference in length between the two constituents was held fairly constant, but the stimuli from Experiment 1 were modified so that the theme NP was always longer than the goal NP (an average difference of 3.6 words). This change was necessary to encourage speakers to produce the double-object construction, since pilot data revealed a preference for the prepositional construction in this task. For the same reason, 4 of the filler items were designed to prime the double object construction (two benefactive and two dative, e.g., “The pilot gave the co-pilot a series of garbled commands during take-off.”) A programming error resulted in all of the prime fillers occurring immediately before ambiguous stimulus items. However, an analysis of items in the ambiguous condition shows that the prime had no effect on the rate of goal-early structures for immediately following items ($\chi^2 < 0.8$, $p > 0.3$), in accord with recent findings that syntactic constructions remain active for a substantial period of time after they are primed (Bock & Griffin, 2000).

The advantage of using a task like this was that it allowed us to control the length and ambiguity of the relevant constituents, but left the choice of construction up to the speaker. The presence of the addressee, who had to answer specific questions, also turned it into more than

just a memorization task. If ambiguity avoidance affects language production at all, it may tend to occur most when speakers care about the communicative value of their speech. We therefore took two measures to encourage participants to think of this task as a communicative enterprise. First, participants were given ten practice items, during which time the speaker was encouraged to consult the listener's response sheet after each item, to see whether they had achieved the correct answer. These items were also designed to be fairly difficult, so speakers would realize it was important to remember the constituents accurately and produce comprehensible utterances. During the rest of the experiment the speaker did not consult the answer sheet, but the listener was also asked to call out the letter of the answer ("a" or "b") to the experimenter. This had the purpose of continuously reminding participants that it was important to get the right answer, thereby emphasizing the communicative nature of the task.

Participants

28 pairs of native speakers of English participated in the experiment in exchange for \$10 each.

Transcription and Coding

Each session was tape-recorded, transcribed and coded along several dimensions. First we coded whether the speaker used the theme-early prepositional construction, or one of two goal-early constructions: the double-object or shifted prepositional (see table 3). Second, we noted which verb was used in each response, and counted the actual number of words in each constituent.

[INSERT TABLE 3 ABOUT HERE]

Responses were excluded if the meaning of the constituents was not preserved, or if the response did not use a form that could participate in the Dative alternation. This resulted in a total of 317 items (out of 448). Of these, we only included those items that preserved the ambiguity manipulation (86%, $N = 272$), excluding those cases where the participant changed the wording in such a way as to add or remove an ambiguity. Notably, changes in wording resulted in adding a potential ambiguity slightly more often (16%) than they resulted in removing an ambiguity (13%).

The entire data set was coded by one experimenter using the above schema, and checked by a second experimenter. The reliability of the coding schema was established by having a third experimenter code a sub-set of 92 randomly chosen items. 83 of these items were coded identically (90%). An additional 4 items included a small difference that was not critical to the final analysis; if we count these, the inter-rater reliability was 95%.

Results

We examined the results by asking what percentage of responses were in the goal-early order, in each ambiguity condition. In contrast with the comprehension results in Experiment 1, there was little effect of the ambiguity manipulation. Furthermore, the numerical results went in the opposite direction of that predicted by an ambiguity avoidance strategy. The percentage of goal-early constructions produced in the unambiguous condition was 80% ($SE = .05$), and in the ambiguous condition was 67% ($SE = .05$). That is, participants were slightly more likely to use

the goal-early order when there was no potential ambiguity than when there was one. However, this difference was not statistically robust.

These differences were submitted to ANOVAs with both participants and items as random variables. One participant was missing data for one condition, and was therefore excluded from the analysis. The main effect of ambiguity was only marginal in the participants analysis, and was unreliable in the items analysis $F(1,26) = 3.7$ $p = .065$, $F(1,15) = 2.82$, $p = .114$).

Thus, ambiguity had little effect on ordering preferences, and in the opposite of the expected direction. Is this effect just noise, or is there something else that might explain the observed ordering preferences? We conducted two additional analyses on these data to investigate the possibility that these ordering preferences come not from the ambiguity of the item, but other differences between the responses in each condition. Since this is a production experiment, participants introduced variation into their responses that may have affected ordering. We therefore investigated whether the ambiguity effect held in the presence of two factors that are known to influence choices in constituent ordering, 1) individual verb biases and 2) syntactic weight. These analyses also tested whether our study was sensitive enough to measure variations in ordering preferences, given that the manipulated variable (ambiguity) had little effect.

Lexical Bias

Dative utterances and other sentences with the multiple postverbal constituents are known to exhibit tendencies for some verbs to occur in one construction more than the other (Stallings et al, 1998; Wasow, 1997). We assessed the lexical bias associated with each verb

used in our study with an analysis of the New York Times corpus. We searched for as many utterances as possible (up to about 100) containing the correct verb in either the double object or prepositional construction (i.e., Verb+NP+PP or Verb+NP+NP). Since the shifted forms are relatively rare for these verbs, they were not included in the corpus analysis. The number of tokens found for each verb ranged from 5 to 102 (average 55). These verbs were divided into “high” and “low” bias verbs so that approximately half the experimental data fell into each category. The “high bias” items were those that occurred in the double-object construction 42% of the time or more. There were also two verbs that were only used once, in Experiment 3; these were not included in the corpus analysis and were grouped with the “low bias” verbs.

Table 4 shows the categorization of "high" and "low" bias verbs, where “high” corresponds to a high rate of goal-early (double-object) constructions. Note that there are more verbs than those used in the stimulus items; this is because speakers sometimes replaced lexical items with roughly synonymous ones.

[INSERT TABLE 4 ABOUT HERE]

We investigated the effect of lexical bias on speaker response for each construction by crossing lexical bias with the ambiguity manipulation and calculating the participant mean for each condition. Six participants were missing data in one or more conditions and were thus excluded from the analysis. Analyses on the item means were impractical because participants often replaced lexical items with roughly synonymous ones, such that 44% of the items occurred with more than one verb.

The results (Figure 1) show that more goal-early constructions were used for high-bias than low-bias Dative items ($F(1,21) = 8.97, p = .007$). Lexical bias also interacted with ambiguity ($F(1,21) = 5.65, p = .027$), such that an ambiguity difference only occurred for low-bias verbs. The main effect of ambiguity was not reliable ($F(1,21) = 2.338, p = .141$).

[INSERT FIGURE 1 ABOUT HERE]

These results suggest that if there is any reverse ambiguity effect, it is small compared to lexical bias, and is more likely to show up for verbs that are not already biased toward the goal-early construction.

Weight

The most widely recognized influence on ordering preferences is syntactic weight, such that long, complex phrases tend to follow shorter, less complex ones (e.g., Behagel, 1909/10; Hawkins, 1994; Wasow, 1997). Several criteria have been reported for measuring the weight of a constituent, including the number of words, number of nodes, or number of phrasal nodes. All of these measurements correlate highly with ordering preferences, and with each other (Wasow, 1997). It has also been suggested that the relative weight of two constituents is more important than the absolute weight of either (Hawkins, 1994; Wasow, 1997).

In this experiment we did not manipulate the weight of the phrases in the stimulus material. Weight was controlled in terms of the number of words in each constituent: in all cases the theme was a few words longer than the goal. However, participant variation in the production of the phrases resulted in length differences ranging from 0 to 7 words.

We categorized the relative length difference between constituents as either “large” (a difference of 4 words or more), or “small” (0-3 words). This factor was crossed with the ambiguity manipulation, and the participant means calculated for each cell. Six participants were excluded because they were missing data in one or more conditions. As with the lexical bias analysis, item means could not be calculated because of variation across participants in the number of words used for each constituent in a given item. The results are shown in Figure 2.

[INSERT FIGURE 2 ABOUT HERE]

The data show a numerical trend towards an interaction, in that the contrast between the ambiguity conditions was greatest for items with a small weight difference. However, the ANOVA revealed only a marginal main effect of weight $F(1,21) = 3.01, p = .097$. No other effects were reliable (p 's > .2)

Like the lexical bias analysis, these results suggest that the reverse ambiguity effect was smaller than the effect of weight, which was not a manipulated variable.

Discussion

Contrary to expectations, the results from experiment 2 showed that speakers do not use constituent ordering to avoid temporary PP-attachment ambiguities. If they did, we should have seen a higher rate of goal-early constructions in the ambiguous condition. Instead, participants showed a weak tendency to use these constructions more in the unambiguous construction. However, this trend was small compared with the effects of lexical bias and syntactic weight.

Furthermore, it tended to occur when weight differences were small and did not strongly constrain the choice of order.

These results leave us with two questions. First, why do speakers not use constituent ordering to avoid ambiguities? We will turn to this question in detail in the general discussion. Second, is the weak effect of ambiguity truly reflecting a tendency for speaker to use theme-early forms more often when doing so creates a PP-attachment ambiguity? The marginality of this effect opens the possibility that it is spurious. We therefore used a second on-line experiment to further explore the effect of ambiguity avoidance on constituent ordering. This also allowed us to replicate the analyses with weight and lexical bias, which were conducted post-hoc in Experiment 2.

EXPERIMENT 3

Experiment 3 provided a more stringent test of whether speakers ever use an ambiguity avoidance strategy, by encouraging speakers to believe that ambiguity avoidance both matters and is necessary. We changed the instructions to highlight the importance of designing the utterances to make them easy for the addressee to understand. It may be the case that speakers only bother to avoid ambiguity when they are aware (either explicitly or implicitly) that the utterance may be difficult to understand, and when they care to make it easy to understand. We reasoned that by practically telling participants to avoid ambiguities, we would provide a more extreme test of whether constituent ordering is ever used as an ambiguity avoidance strategy.

We also attempted to make speakers believe in the importance of producing clear utterances by modifying the practice items. These were the only items on which the speaker

views the addressee's responses and the alternate choices. We therefore made these items even harder than they were in Experiment 2, and embedded several ambiguities. Two had PP-attachment ambiguities (one in a dative construction), and three had different ambiguities (e.g., direct object / S-comp ambiguity). For the dative item, the "wrong" answer was what the addressee would have chosen if they had been garden-pathed by the local PP-attachment ambiguity and never recovered.

The second purpose of this experiment was to investigate lexical bias and weight as planned analyses, and explicitly manipulate lexical bias.

Methods and Materials

The same methods and materials from Experiment 2 were used, with a few modifications. The critical difference was in the instructions, which encouraged the speakers to provide informative, easy-to-understand answers to the listeners. For example, speakers were told "You should also know that the design of the experiment sometimes leads people to produce sentences that are awkward or hard to follow. This may be unavoidable, but you should do your best to make it easy for your partner to understand your sentences." Immediately before the experiment began, they were reminded to make their response easy to follow.

As in Experiment 2, the communicative nature of the task was also emphasized by having the listener call out the answer to the experimenter, as a way of reminding the participants that they should be aiming to get the right answer. One experimenter asked participants to say only the letter for the answers, and another asked them to say the entire sentence. This difference did not affect the speakers' choice of construction .

The same sixteen dative items from Experiment 2 were used, eight with high-bias verbs and eight with low-bias verbs. As before, each item had two variants, one with a potential ambiguity and one with no potential ambiguity (see Table 2). These items were pseudorandomly combined with sixteen heavy-NP-shift items (also in high and low bias conditions), and 48 fillers. They were organized into two new lists (with forward and backward versions). Each item occurred only once per list, in either the ambiguous or unambiguous condition, and each list was designed to contain equal numbers of items in the four conditions that resulted from the cross of ambiguity and lexical bias. However, one item had to be excluded from the analyses because it accidentally appeared in the ambiguous condition in both lists. The 48 fillers included 4 primes, as in Experiment 2, but this time each prime occurred before another filler item, so no stimulus item appeared immediately after a prime.

Participants

36 pairs of native speakers of English from the Stanford community participated in the experiment in exchange for \$10 each.

Results

The responses were transcribed and coded by the same criteria as in Experiment 2. 400 out of 540 items met the criteria of preserving meaning and structure (74%). Of these, 60 were excluded because the participant changed the wording in such a way that it added or removed the potential ambiguity. This occurred more often for the ambiguous items (19%), but still occurred quite often for unambiguous items (11%; difference: $\chi^2 = 4.01$, $p < 0.05$).

The results (Figure 3) revealed an even clearer reverse ambiguity effect than in Experiment 2, as well as a strong effect of lexical bias. An ANOVA was conducted on participant means, excluding the four participants with data missing from one condition. The ANOVA showed a main effect of Ambiguity ($F(1,31) = 11.83, p < .005$), a main effect of Lexical Bias ($F(1,31) = 11.49, p < 0.005$), and no interaction between Ambiguity and Lexical Bias ($F(1,31) = 0.94, p > .3$). As in Experiment 2, the items analysis was not possible because of participant variation in the verb used for a given item.

[INSERT FIGURE 3 ABOUT HERE]

Again we conducted a second analysis, crossing ambiguity with weight (Figure 4). The results revealed strong effects of both weight and ambiguity. An ANOVA over participant means, excluding the 8 participants with missing cells, showed a main effect of Ambiguity ($F(1,27) = 13.18, p = 0.001$), a main effect of Weight ($F(1,27) = 15.33, p < .001$), and no interaction between the two ($F(1,27) = .562, p > .4$).

[INSERT FIGURE 4 ABOUT HERE]

The results for Experiment 3 show a clearer and larger effect of weight than those for Experiment 2, even though this factor was not manipulated in either experiment. This is likely to stem from the change in instructions, which essentially encouraged participants to deviate from the given stimuli when necessary. Even though the experiments used the same stimuli,

Experiment 2 produced weight differences ranging from 0 to 7 words, while the range in Experiment 3 was -4 to 10 words.

Prosodic Analyses

Experiment 3 provided strong evidence that speakers do not use constituent ordering to avoid ambiguities. Is this because other mechanisms, like prosodic cues, helped disambiguate their utterances? Some studies have found that the location and intensity of intonational breaks facilitate certain interpretations of PP-attachment ambiguities (e.g., Shafer, 1997; Snedeker & Trueswell, 2003; Warren et al., 2000; Watson & Gibson, in press, a). For example, in a sentence like “Tickle the frog with the flower”, a break after “tickle” signals NP-attachment of the phrase “with the flower”, while a break after “frog” signals VP attachment (Snedeker & Trueswell, 2003). This suggests that if prosodic cues were strong enough in our data, there would be little reason for speakers to avoid the prepositional order.

Several theoretical approaches explain the disambiguating role of prosodic marking (e.g., Carlson, Clifton, and Frazier, 2001; Price et al., 1991; Shafer, 1997; Watson and Gibson, in press, a). Here we focus on the predictions of Watson and Gibson’s (in press, a) Anti-Attachment Hypothesis, which states that an intonational boundary immediately following a lexical head indicates that no further attachments will be made to that head. In our data, the PPs were always attached to the noun, so producing a break before the ambiguous PP would be a false cue that it should not be attached to the direct object noun, inhibiting comprehension. We therefore predicted few breaks before the embedded PP. However, there is no felicitous way for speakers to indicate the correct structure prosodically. Unlike most research on this issue, which has investigated sentences with single prepositional phrases, our data always contained two

prepositional phrases. The Anti-Attachment hypothesis predicts that a break after the verb signals no further attachments to the verb – yet the second PP does in fact attach to the verb. Therefore, we might only expect a large number of breaks after the verb if speakers were specifically aiming to avoid the PP-attachment ambiguity, at the cost of processing difficulty for the second PP.

We asked whether speakers disambiguated their prepositional constructions using either of these cues, by analyzing the incidence of breaks both after the verb, and before the embedded phrase. These were compared to a third location, before the goal prepositional phrase (see Table 5). Breaks before the goal PP would come too late to disambiguate the embedded PP, but would be generally consistent with the structure of the utterance.

[INSERT TABLE 5 ABOUT HERE]

We measured the breaks in speakers' responses both acoustically and phonologically. First, we used the phonetics software Praat (Boersma & Weenink, 2002) to measure the duration of the pause between words at each of those locations. Second, a research assistant who was naïve to the experimental hypotheses identified the location and strength of all breaks, using a 7-point scale based on the ToBI system (Silverman, et al., 1992). All breaks greater than expected between two words were identified, and the strength of the break was coded on the scale 4, 3.5, 3, 2.5, 2, and 1.5. The absence of a break was coded as a 1 (in the ToBI coding system this would correspond to a break of 1 or 0). Codings of levels 2.5-4 correspond to intermediate and full intonational breaks; these were categorized together as “break”, and codings of 1-2 were categorized as “no break”. The same data were also analyzed as a continuous variable.

One entire item (#6) was excluded from analysis because the unambiguous condition used pre-head modification, making it impossible to compare the prosody across conditions (“tenth-grade level math workbooks” vs. “workbooks for tenth graders”). Two individual responses were excluded for the same reason, and one participant was excluded because he was a native speaker of a non-American version of English. Additional responses were excluded because the tape was distorted or unavailable. This left 92 items in the analysis, from 28 speakers and 14 items. 11 speakers and 3 items only had data in one ambiguity condition or another, and had to be excluded from the statistical analyses. Although this left only 17 participants and 11 items in the analysis, they displayed the same patterns as the overall participant means. Here we report the average participant means for those participants with data in both conditions.

The main result was that speakers produced few positive prosodic cues to help disambiguate the ambiguous PP. Breaks after the verb, which would have facilitated NP-attachment of the embedded PP, occurred only 18% of the time. As mentioned above, this low rate of breaks was expected, since the later PP does attach to the verb, even though such a break would have effectively disambiguated the attachment of the embedded PP. By comparison, breaks before the goal PP were more common, 34%. The best cue speakers produced was not a positive cue, but the lack of a misleading breaks before the embedded phrase. Misleading breaks occurred only 21% of the time, and more than half of those breaks (9 out of 15) co-occurred with a break at one of the other two locations, which may have decreased their negative impact.

We also asked a second question, whether speakers were less likely to produce an unhelpful break when the utterance was pragmatically ambiguous than when it wasn't. This second question speaks to the debate on whether prosodic cues to structure are produced as an

ambiguity avoidance strategy, as Snedeker et al. (2000) found for global PP-attachment ambiguities. Alternatively, prosodic patterns may be produced automatically, based on the syntactic structure of the utterance, rather than to avoid ambiguity per se (Kraljic & Brennan, 2003; Warren et al., 2000).

The results initially seem consistent with the strategic avoidance of misleading cues. Speakers were less likely to produce a break before the embedded phrase in the ambiguous condition (11%), compared with the unambiguous condition (31%). This difference only reached significance in the participants analysis ($F(1,16) = 8.13, p < .05$; $F(1,10) = 2.874, p = .121$). However, an analysis of the same data as a continuous variable was statistically robust, where the average break strength was greater in the unambiguous condition (average 1.89) than the ambiguous condition (1.32) ($F(1,16) = 11.75, p < 0.005$; $F(1,10) = 5.00, p < .05$). The duration of pauses immediately before the embedded phrase show a similar pattern of results (ambiguous condition: 113 msec; unambiguous condition: 357 msec; $F(1,16) = 4.47, p = .051$; $F(1,10) = 2.135, p = 0.175$).

However, the structural view of the production of prosody could also potentially account for these results. The ambiguous responses always contained a “to”-PP, which might be considered an argument of many direct object nouns in our stimuli (see Schütze, 1995, for some diagnostics for argumenthood). The prepositional phrases in the unambiguous responses (“about...”, “before...”, etc.) may not all be arguments, and the verb phrases that occurred in some items (e.g. “praising Lincoln”) are certainly not arguments of the head noun. Some production models of prosody predict a dispreference for intonational breaks between a head and its argument (e.g., F. Ferreira, 1988), which would account for the difference between conditions here.

Whatever the reason for the intonational phrasing, the results from these analyses show that speakers tend to produce a prosodic pattern that is generally consistent with the intended structure. They produced few breaks after the verb or before the embedded constituent, which would falsely signal no further attachments. Breaks before the goal NP, which are consistent with the structure, were more common.

However, it is not clear that speakers produced a prosody that would disambiguate the embedded PP, either intentionally or unintentionally. The most positive cue to NP-attachment, a break after the verb, was infelicitous for other reasons, and therefore occurred rarely. The only helpful cue was a “null” cue – that is, the lack of a misleading break. This prosodic structure probably did facilitate the recovery of the NP-attached structure, at least partially. But its effect would be strongest if listeners did not attribute the lack of a break to one of the many other determinants of prosodic phrasing. For example, speakers are more likely to place prosodic breaks immediately preceding or following complex constituents (Watson and Gibson, in press, b). In our stimuli, these constituents (e.g., “Grant’s letters”, or “to Lincoln”) may not have been complex enough to warrant the placement of a break, in which case the lack of such a cue would not be as informative.

Thus, the prosodic analyses do not provide compelling evidence that the responses in our experiment were disambiguated sufficiently to make it unnecessary to use other mechanisms for disambiguation. While prosody may partially facilitate parsing, even in the best of cases it poses a weak and partial cue to structure. By comparison, constituent ordering offers a way of definitively marking the intended structure, if speakers wished to disambiguate the utterance.

Discussion

In this experiment, where the instructions stressed the importance of producing easy-to-follow utterances, again speakers did not use constituent ordering as an ambiguity avoidance mechanism. Prosodic analyses of responses revealed that speakers tended not to provide positive cues that would disambiguate the utterance. Moreover, even though speakers were less likely to produce a misleading break in the ambiguous condition, we cannot conclude that speakers use prosody as a mechanism for disambiguation, since this difference could be attributed to differences between conditions in the argumenthood of the embedded phrases. Either way, prosodic cues in these utterances would at best provide only weak cues about the attachment of the embedded PP, in comparison with the potentially stronger cue of constituent ordering. These findings together raise the question of why speakers do not do more to help listeners avoid a garden path, which we will discuss in the next section.

The more surprising result from Experiment 3 was a more stable reverse ambiguity effect than in Experiment 2. The reverse ambiguity effect occurred alongside the more expected effects of lexical bias and weight, and did not interact with either. Thus, it seems this is a real (albeit small) effect.

Why, then, do speakers prefer the theme-early orders for the ambiguous utterances? We can offer two speculations. One possibility is that the theme-early order is chosen precisely because the production system is sensitive, at some level, to the comprehension difficulty of these utterances. The presence of two goal arguments in the ambiguous condition may be a flag to speakers that the utterance will be complex and difficult to understand. This difficulty may lead speakers to choose the order that marks the goal argument with an explicit morphological marker (namely, the preposition “to”), rather than indirectly by means of word order. This

construction also offers the more frequent mapping between thematic roles and grammatical functions throughout the language. Theme arguments tend to be mapped to the direct object in many transitive verbs, which may make it easier to comprehend dative sentences that follow this same mapping. Both the explicit marking of the goal and using a construction with the more frequent mapping between thematic roles and grammatical functions should make the utterance easier to understand. The ambiguous conditions may also be simply more difficult for the speaker to plan and produce, perhaps because of the overlap in thematic representation of the two goal constituents,¹ leading speakers to resort to the more common thematic role/grammatical function mapping as a way of facilitating the production process.

This hypothesis can be tested by taking a further look at the data, which are reported above in terms of the theme-early (prepositional) vs. goal-early (double-object + shifted-prepositional). If speakers tend to prefer constructions that explicitly mark the goal argument when the utterance is harder to follow, then we would expect that the goal-early items would have a higher proportion of shifted-PP items in the ambiguous than the unambiguous conditions. This turns out to be the case for both experiments (Experiment 2: $F(1,24) = 6.74, p < .05$; Experiment 3, $F(1,29) = 3.84, p = .059$).² The participant means in each category are shown in Table 6.

[INSERT TABLE 6 ABOUT HERE]

¹ We thank Duane Watson for this observation.

² Participants who did not produce any goal-early constructions in one condition or the other were excluded from this comparison.

Not only were there more prepositional responses in the ambiguous conditions, there was also an increase in the percentage of prepositional responses overall in Experiment 3, where the instructions emphasized the need to be clear (101/272 in Experiment 3, 152/340 in Experiment 2, $\chi^2(1) = 3.57, p = .059$). This suggests that the instructions may have resulted in a strategy of explicitly marking thematic roles with the prepositional construction. A similar effect occurred in V. Ferreira and Dell's (2000) Experiment 6: participants who were encouraged to focus on the task as communicative included more optional *that*'s; but contrary to the current findings, *that*-usage did not vary according to the ambiguity of the sentence. Similarly, our findings suggest that speakers may engage in general strategies to clarify their speech when they are paying attention to the need to be clear. However, the choice of the prepositional construction falls short of addressing the actual needs of comprehenders, who clearly dispreferred this structure in Experiment 1.

A second possible explanation for the reverse ambiguity effect is priming from the NP-internal prepositional phrase.³ Recall that the stimulus items did not occur in any of the target orders, but they did include the full target constituents, for example “A museum received Grant's letters to Lincoln from the foundation.” All the ambiguous items contained a PP within the NP (e.g., “to Lincoln”), which could have primed the prepositional frame in the participant's response. It is also notable that the preposition itself is the same as the one that would be needed for the prepositional frame (e.g., “gave letters **to** Lincoln **to** the museum”). Although syntactic priming is not dependent on the particular lexical items used (Bock, 1989), it is also possible that the lexical item may have become primed, thus priming its use in the response. This is consistent with findings that syntactic priming is stronger when lexical items are shared between

prime and target (e.g., Branigan, Pickering, & Cleland., 2000; Pickering & Branigan, 1998) and can even occur when the prime contains the same function word but with a different syntactic structure (Bencini, Bock, & Goldberg, 2002; but see V. Ferreira, 2003 for a different finding). By contrast, the unambiguous items did not uniformly contain an NP-internal PP. Five items removed the ambiguity by rephrasing the NP, e.g. “A museum received Grant's letters praising Lincoln from the foundation.” The other 11 items contained a PP, but using a different preposition. Thus, the unambiguous items as a group may not have primed the prepositional frame as strongly, and certainly did not prime the preposition “to” at all.

However, this explanation is unlikely to fully account for the reverse ambiguity effect. If syntactic priming were responsible, we would expect a greater use of the theme-early prepositional structure in those items in which the unambiguous condition contained an NP-internal PP, compared with those items where it did not. That is, we might have expected that the reverse ambiguity effect would be carried by the items where the ambiguous version had an NP-internal PP, and the unambiguous version did not. But as Table 7 shows, neither experiment showed a large difference in the expected direction, and in fact in Experiment 2 the reverse ambiguity effect was substantially larger for those items that did have a preposition in the unambiguous condition.

[INSERT TABLE 7 ABOUT HERE]

Regardless of the reason for the reverse effect, these data clearly show that constituent ordering does not provide a natural mechanism for avoiding ambiguities during production.

³ We gratefully acknowledge Gary Dell for this suggestion (personal communication).

Even when speakers were encouraged to think about the need to produce clear and easy-to-understand utterances, they did not choose word orders that would avoid PP-attachment ambiguities.

GENERAL DISCUSSION

The results from the two production experiments showed that speakers do not choose goal-early orders for dative utterances more when that order would avoid an ambiguity. Instead, ordering choices are consistently influenced by weight and lexical bias. These results stand in contrast to the results from Experiment 1, which showed that comprehenders preferred the goal-early order more when it avoided an ambiguity. This strongly suggests that if speakers avoid ambiguities at all during language production, they do not use constituent ordering to do so.

This raised the question whether speakers disambiguate in some other way. We considered one possible mechanism: intonational phrasing. We found that indeed speakers produced 20% fewer misleading breaks in the ambiguous condition than in the unambiguous condition. However, it is not clear that this difference was caused by a desire to avoid ambiguity, since an argumenthood difference between conditions might account for the same pattern. In sum, we found very little evidence that speakers strategically avoided local PP-attachment ambiguities, either with constituent ordering or prosody.

On one hand it seems surprising that speakers do not bother to avoid attachment ambiguities, since such ambiguities are known to impair language comprehension. Given that variation occurs in both ordering and intonational phrasing, it would seem natural for speakers to capitalize on one mechanism or the other in order to make themselves better understood.

However, our findings are not surprising if we consider the cognitive machinery that would be required, and other cues that might attenuate the processing difficulty of attachment ambiguities.

First of all, the process of ambiguity identification alone may be unwieldy. Language production involves generating an utterance from a non-linguistic message. The message is never ambiguous to the speaker; the only way to identify the ambiguity is to consider how someone else would interpret the message in the current context. This would require passing the planned utterance through the comprehension system, while ignoring the known intended meaning. The production system would have to be sensitive to the degree of temporary parsing difficulty associated with an ambiguous prepositional phrase, and use that information to drive decisions about ordering and prosody. It is not clear that the language production system is built to handle this kind of task. Although language production clearly involves monitoring at some level (see Postma, 2000, for a review), the clearest application of these monitors is to the process of identifying and correcting errors. Ambiguities are not errors per se, and may require more sophisticated machinery for identifying them. Indeed, speakers sometimes underestimate the difficulty of interpreting global PP-attachment ambiguities (Keysar & Henly, 2002).

Apart from issues of feasibility, speakers may not be motivated to avoid attachment ambiguities, especially if they are only local. While garden paths can be disruptive, most of the time language is produced in a rich context, and comprehenders are good at using multiple sources of information to help guide ambiguity resolution (e.g., MacDonald, Pearlmutter & Sedienberg, 1994; Tanenhaus & Trueswell, 1995). Even in written language, where speakers can record phrases on paper and reorder them later if necessary, it is not difficult to find examples containing exactly the kind of ambiguity studied here. Examples like (3) use the

ambiguous prepositional dative order, even though the embedded PP provides a plausible goal and could easily cause a garden-path.

(3) But soon another Florentine Dominican, Niccolo Lorini, submitted a copy of Galileo's now widely read letter to Castelli -- to an inquisitor general in Rome.
(Dava Sobel, Galileo's daughter, p. 67)

In addition to the above difficulties, the use of constituent ordering for ambiguity avoidance would pose particular cognitive demands. Speakers would need to formulate at least the ambiguous constituent, which is likely to be complex, if not both constituents. The constituent(s) would then have to be maintained in memory while monitoring for an attachment ambiguity and choosing the order in which to produce them. This is likely to be especially difficult, because it requires more pre-planning than speakers normally engage in. There is reason to believe that ordering decisions are often made before completing constituent formulation (Wasow, 1997). Speakers often begin speaking before the entire utterance is formulated, given the incremental nature of language production (Clark & Wasow, 1998; V. Ferreira, 1996, Kempen & Hoenkamp, 1987). Instead, ordering decisions and grammatical function assignments are influenced by lexical and conceptual availability (Bock, 1982; 1986; Bock & Warren, 1985; Bock & Irwin, 1980; V. Ferreira & Dell, 2000), which means they often occur early, before the ambiguous portion of the constituent is even formulated.

By comparison, the choice of where to place intonational breaks may be more cognitively tractable as an ambiguity avoidance mechanism, since the choice of where to place a break does not have to occur until shortly before the break location. However, prosody is not an airtight

indicator of the intended structure, since any given choice of intonational phrasing could be consistent with multiple syntactic structures, in addition to being influenced by processing pressures. It may also be particularly hard to signal NP attachment prosodically when a temporarily ambiguous PP is followed by a VP-attached PP. The clearest cue against VP attachment is a break after the verb, but that would be inconsistent with the later need to attach the goal PP to the verb (Watson & Gibson, in press, a). This does not mean that speakers never use prosody to disambiguate, but rather that it may be difficult to do so with structures like the one investigated here.

In sum, these experiments provide clear evidence that speakers do not use constituent ordering to avoid ambiguities. Prosodic analyses revealed that speakers indeed produced ambiguous utterances when they chose the theme-early form, in that the pattern of intonational breaks was not enough to consistently disambiguate the utterance.

But these results should not be taken to mean that language production is not influenced by the desire to be clear at any level. For example, speakers may consider the ambiguity of referential expressions before producing them (Brown-Schmidt, Campana, & Tanenhaus, 2003; Nadig and Sedivy, 2002), at least when time constraints permit (Horton & Keysar, 1996). In addition, the results above suggest that speakers may even consider listener needs when choosing syntactic constructions. Although we originally set out to study whether constituent ordering was driven by the desire to avoid ambiguity, ordering is partially confounded with the choice of a syntactic construction (double-object or prepositional). One explanation for the small reverse-ambiguity effect is that speakers want to explicitly mark the goal argument. This suggests that while linear ordering is not directly influenced by communicative pressures, other syntactic choices may be, and these choices in turn affect ordering. But this reaction to the desire to be

clear is different from avoiding ambiguities. It does not involve the anticipation of an ambiguity, but rather the simple addition of a function word to explicitly mark thematic role assignment.

This process is unlikely to consume memory resources, in comparison to the process of predicting ambiguities and re-ordering entire constituents. Furthermore, this choice may be made as a result of a general desire to “be clear”, as opposed to avoid structures that are specifically awkward for comprehenders. Experiment 1 established that comprehenders prefer the goal-early (double-object) order in the ambiguous condition, but the speakers in Experiments 2 and 3 preferentially chose theme-early orders in this condition.

A second way in which participants may have attempted to clarify the ambiguous stimuli was through changes in the wording of the constituents. These changes sometimes added or removed the potential ambiguity that had been in the stimulus. For example, when the stimulus was “The secretary was brought the memo to the president by the courier,” some responses avoided the ambiguity through wording, for example “Brought the memo for the president to the secretary.” However, sometimes a change in wording actually added an ambiguity. For example, when the target response (in theme-early form) was “The executive committee faxed the tribute for the outgoing director to everyone,” one participant responded “They faxed a tribute to a top level individual to everyone.” In Experiment 2, both types of wording changes occurred about equally, but in Experiment 3, participants were more likely to remove an ambiguity than to add one. This may indicate a greater desire to avoid ambiguities in Experiment 3, where the need for clarity was emphasized. However, the fact that ambiguity additions occurred as often as they did (16% in Experiment 2, 11% in Experiment 3) suggests that participants were not consistently using word choice to avoid ambiguity.

Thus, these studies provide little evidence that production decisions are driven by an ambiguity avoidance strategy. Instead, ordering decisions appear to be driven by other factors, including lexical bias and syntactic weight. To the extent that the needs of the listener are considered, they are addressed with other mechanisms. Speakers may attempt to make difficult utterances clearer by explicitly marking thematic roles, through lexical choice, or by changing the phrasing of the utterance. The answer to our original question is clear: while constituent ordering might have provided a natural mechanism for avoiding ambiguities in production, it appears that speakers do not, or can not, make use of it.

AUTHOR NOTE

Jennifer E. Arnold, Institute for Research in Cognitive Science, University of Pennsylvania and the Dept. of Brain and Cognitive Sciences, University of Rochester. Thomas Wasow and Ash Asudeh, Dept. of Linguistics, Stanford University. Pete Alrenga, Dept. of Linguistics: Stanford University (now at the Dept. of Linguistics, University of California, Santa Cruz).

Ash Asudeh was partially supported by a SSHRC doctoral fellowship 752-98-0424. We owe many thanks to Nina Kim for her help designing Experiment 1 and collecting the data, to Kerry Danahy and Cheryl Andrada for their help in collecting and coding the data for Experiment 3, to Matt Flannery for his help with the lexical bias corpus analysis, and to Jill Thorson for her help with the prosodic analysis. We thank the following people for their comments and discussion: Lera Boroditsky, Herb Clark, Gary Dell, Ted Gibson, Zenzi Griffin, Michael Tanenhaus, Duane Watson, and the language community at the University of Pennsylvania. An earlier version of this paper was presented at AMLaP 2000 in Leiden, and a brief report of these data is presented in Wasow, T., and Arnold, J. E. (2003). Postverbal constituent ordering in English. In G. Rohdenburg & B. Mondorf (Eds), *Determinants of Grammatical Variation in English*. The Hague: Mouton, 120-154.

REFERENCES

- Altmann, G. & Steedman, M. (1988). Interaction with Context During Human Sentence Processing. Cognition **30**, 191-238.
- Arnold, J. E., Wasow, T., Ginstrom, R., & Losongco, A. (2000). Heaviness vs. Newness: the effects of structural complexity and discourse status on constituent ordering. Language **76**(1), 28-55.
- Behaghel, O. (1909/10). Beziehungen zwischen Umfang und Reihenfolge von Satzgliedern. Insogermanische Forschungen **25**, 110-42.
- Bencini, G., Bock, J. K. & Goldberg, A. (2002). How abstract is grammar? Evidence from structural priming in language production. Poster session presented at the CUNY conference on human sentence processing, New York, NY.
- Bock, J. K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. Psychological Review, **89**(1), 1-47.
- Bock, J. K. (1986). Meaning, Sound, and Syntax: Lexical Priming in Sentence Production. Journal of Experimental Psychology: Learning, Memory, and Cognition **12**(4), 575-586.
- Bock, J. K. (1989). Closed-class immanence in sentence production. Cognition **31**(1), 163-186.
- Bock, J. K. & Irwin, D. E. (1980). Syntactic effects of information availability in sentence production. Journal of Verbal Learning and Verbal Behavior **19**, 467-484.

- Bock, K. & Griffin, Z. M. (2000). The Persistence of Structural Priming: Transient Activation or Implicit Learning. Journal of Experimental Psychology: General 129(2), 177-192.
- Bock, J. K., & Warren, R. K. Conceptual accessibility and syntactic structure in sentence formulation. *Cognition*, 21(1), 47-67.
- Boersma, P. & Weenick, D.(2002). Praat. [Computer software]. Amsterdam, The Netherlands: Institute of Phonetic Sciences, University of Amsterdam.
- Boland, J. E. & Boehm-Jernigan. H. (1998). Lexical constraints and prepositional phrase attachment. Journal of Memory & Language 39(4), 684-719.
- Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co-ordination in dialogue. Cognition 75, B13-25.
- Brown-Schmidt, S., Campana, E., & Tanenhaus, M. K. (In press). Real-time reference resolution by naïve participants during a task-based unscripted conversation. In J. C. Trueswell & M. K. Tanenhaus (Eds.), *World-situated language processing: Bridging the language as product and language as action traditions*. Boston, MA: MIT press.
- Carlson, K., Clifton, Jr., C., and Frazier, L. (2001). Prosodic Boundaries in Adjunct Attachment. Journal of Memory and Language 45, 58-81.
- Christianson K., Hollingworth, A., Halliwell, J. F., & Ferreira, F., 2001. Thematic roles assigned along the garden path linger. *Cognitive Psychology*, 42(4), 368-407.
- Clark, H. H. & Krych, M. A. (accepted pending revision). Speaking while monitoring addressees for understanding. Journal of Memory and Language.

- Clark, H. H. & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. Cognition 22, 1-39.
- Clark, H. H. & Wasow, T. (1998). Repeating words in spontaneous speech. Cognitive Psychology 37(3), 201-242.
- Ferreira, F. (1988). Planning and timing in sentence production: The syntax-to-phonology conversion. Unpublished doctoral dissertation, University of Massachusetts, Amherst, MA.
- Ferreira, V. S. (2003). The persistence of optional complementizer production: Why saying “that” is not saying “that” at all. Journal of Memory and Language 48, 379-398.
- Ferreira, V. S. (1996). Is It Better to Give Than to Donate? Syntactic Flexibility in Language Production. Journal of Memory and Language 35(5), 724-755.
- Ferreira, V.S., & Dell, G.S. (2000). Effect of ambiguity and lexical availability on syntactic and lexical production. Cognitive Psychology, 40, 296-340
- Frazier, L. (1978). On comprehending sentences: Syntactic parsing strategies. Unpublished doctoral dissertation, . University of Connecticut, Storrs, CT:
- Frazier, L. (1985). Syntactic Complexity. In D. Dowty, L. Karttunen, and A. M. Zwicky (Eds.), Natural Language Parsing (pp. 129-189). Cambridge, England: Cambridge University Press.
- Hankamer, J. (1973). Unacceptable Ambiguity. Linguistic Inquiry 4, 17-68.

Hawkins, J. A. (1994). A performance theory of order and constituency. Cambridge: Cambridge University Press.

Horton, W. S. & Keysar, B. (1996). When do speakers take into account common ground? Cognition 59, 91-117.

Kempen, G. & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. Cognitive Science 11(2), 201-258.

Keysar, B. & Henly, A. S. (2002). Speakers' overestimation of their effectiveness. Psychological Science, 13(3), 207-212.

Kraljic, T., & Brennan, S. (2003). Prosodic disambiguation of syntactic structure: For the speaker or for the addressee? Paper presented at the 16th annual CUNY conference on human sentence processing, Boston, MA: March 27-29, 2003.

MacDonald, M. C., Pearlmutter, N. J. & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. Psychological Review, 1014, 676-703.

Nadig, A. & Sedivy, J. (2000). Finding common ground in children's referential communication. In L. R. Gleitman & A. K. Joshi. Mawah (Eds.), Proceedings of the 22nd Annual Conference of the Cognitive Science Society (p 1046), NJ: Lawrence Erlbaum Associates.

National Archives and Records Administration's Guide to Legal Writing [Electronic Resource].

(2003). Available:

http://www.archives.gov/federal_register/drafting_legal_documents/ambiguity.html

- Pickering, M. J., Branigan, H. P., & McLean, J. F. (2002). Constituent Structure is Formulated in One Stage. Journal of Memory and Language 46, 586-605.
- Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs; Evidence from syntactic priming in language production. Journal of Memory and Language 39, 633-651.
- Postma, A. (2000). Detection of errors during speech production: a review of speech monitoring models. Cognition, 77 97-131.
- Price, P. J., Ostendorf, M., Shattuck-Hufnagel, S., & Fong, C. (1991). The use of prosody in syntactic disambiguation. Journal of the Acoustical Society of America 90, 2956-2970.
- Rayner, K., Carlson, M., and Frazier, L. (1983). The interaction of syntax and semantics during sentence processing. Journal of Verbal Learning and Verbal Behavior, 22 358-374.
- Schober, M. F., & Brennan, S. E. (In press). Processes of interactive spoken discourse: The role of the partner. In A. C. Graesser, M. A. Gernsbacher, & S. R. Goldman (Eds.), Handbook of discourse processes. Hillsdale, NJ: Lawrence Erlbaum.
- Schütze, C. T. (1995). PP Attachment and Argumenthood. In C. T> Schütze, J. G. Ganger, & K. Broihier. (Eds.), Papers on Language Processing and Acquisition (pp. 95-151). MIT Working Papers in Linguistics, 26.
- Shafer, A. (1997). Prosodic parsing: The role of prosody in sentence comprehension. Ph.D. Dissertation. University of Massachusetts, Amherst, MA: GLSA.

- Silverman, K., Beckman, M., Pierrehumbert., J., Ostendorf, M., Wightman, C., Price, P., Hirschberg, J. (1992) ToBI: A standard scheme for labeling prosody. In: Proc. Second International Conference on Spoken Language Processing (ICSLP), Bariff, pp. 867-879.
- Snedeker, J., Gleitman, L., Felberbaum, M., Placa, N., & Trueswell, J. (2000). Prosodic choice: effects of speaker awareness and referential context. Proceedings of the Twenty-second Annual Conference of the Cognitive Science Society. Mahwah, NJ: Lawrence Erlbaum Associates.
- Snedeker, J., & Trueswell, J. (2003). Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. Journal of Memory & Language, 48(1), 103-130.
- Sobel, D. (2000). Galileo's daughter. New York, NY: Penguin Books.
- Spivey-Knowlton, M. J. & Sedivy, J. C. (1995). Resolving attachment ambiguities with multiple constraints. Cognition 55, 227-267.
- Stallings, L. M., MacDonald, M. C. & O'Seaghda, P. G. (1998). Phrasal Ordering Constraints in Sentence Production: Phrase Length and Verb Disposition in Heavy-NP Shift. Journal of Memory and Language 39(3), 392-417.
- Straub, K. A. (1997). The Production of Prosodic Cues and their Role in the Comprehension of Syntactically Ambiguous Sentences. Unpublished doctoral dissertation, Rochester, NY: University of Rochester
- Tanenhaus, M. K. & Trueswell, J. C. (1995). Sentence comprehension. In J. Miller & P. Eimas (eds.). The Handbook of Perception and Cognition: Volume 11, (pp 217-262). Academic Press.

- Warren, P., Schafer, A.J., Speer, S.R., & White, S.D. (2000). Prosodic resolution of prepositional phrase ambiguity in ambiguous and unambiguous situations. *UCLA Working Papers in Phonetics* 99, 5-33.
- Warren, P. (1985). *The Temporal Organisation and Perception of Speech*. Unpublished doctoral dissertation, UK: University of Cambridge.
- Wasow, T. (1997). Remarks on grammatical weight. *Language Variation and Change* 9, 81-105.
- Watson, D. & Gibson, E. (In press, a). Making Sense of the Sense Unit Condition. *Linguistic Inquiry*.
- Watson, D. & Gibson, E. (in press) The relationship between intonational phrasing and syntactic structure in language production. *Language and Cognitive Processes*.
- Yamashita, H. & Chang, F. (2001). "Long before short" preference in the production of a head-final language. *Cognition* 81(2), B45-B55.

APPENDIX: Experiment Materials

Experiment 1. The unambiguous and ambiguous versions are shown here in the prepositional order. Participants were presented with these along with their corresponding versions in the double-object construction.

- 1 The damn computer sent my private message {to/ about} you to everyone on the whole system.
- 2 The teacher read Andy's note {to / about} Terry to the entire English class.
- 3 The foundation gave Grant's letters {to / praising} Lincoln to a museum in Philadelphia.
- 4 The courier brought the memo {to / regarding} the president to the president's own personal secretary.
- 5 The sergeant showed Pat's confession {to / blaming} Sasha to the chief commanding officer.
- 6 The teacher bought {math workbooks for tenth graders / tenth grade level math workbooks} for her advanced ninth grade students.
- 7 The translator told the man's sad testimony {to/ about} the court to the reporters who were in the lobby.
- 8 The New York government sold the bridge {to/ near} Canada to a local investment firm.

Experiments 2/3.

- 1 Everyone received my private message {to / about} you from the damn computer.
- 2 The entire class was read Andy's note {to/ about} Terry by the teacher.
- 3 A museum received Grant's letters {to /praising} Lincoln from the foundation.
- 4 The secretary was brought the memo {to /regarding} the president by the courier.
- 5 The officer was shown Pat's confession {to /blaming} Sasha by the sergeant.
- 6 The ninth graders were bought {math workbooks for tenth graders / tenth grade level math workbooks} by the teacher.

- 7 The reporters heard the man's testimony {to / before} the court from the translator.
- 8 An entrepreneur bought the bridge {to / near} Canada from the New York government.
- 9 The kitchen supervisor was radioed the colonel's orders {to / for} the cook by the junior officer.
- 10 Morgan was shipped the interim report {to / from} the review board by the secretary.
- 11 The new staff were e-mailed the revised instructions {to / about} priority users by the system administrator.
- 12 The reporter was handed the Senator's apology {to / regarding} the panel by the messenger.
- 13 Potential donors were mailed an engraved invitation {to / from} the White House by the Vice President.
- 14 An administrative assistant {to / for} the department head was assigned to every program by the dean.
- 15 Each professor was allotted two new master keys {to / for} the department by the university.
- 16 Everyone was faxed the tribute {to / for} the outgoing director by the executive committee.

TABLES

TABLE 1

Sample Stimuli and Results for Experiment 1

	<u>Stimuli</u>	<u>Results</u>
1. Potential ambiguity:		
	a. The foundation gave a museum in Philadelphia Grant's letters to Lincoln.	53%
	b. The foundation gave Grant's letters to Lincoln to a museum in Philadelphia.	47%
2. No potential ambiguity:		
	a. The foundation gave a museum in Philadelphia Grant's letters praising Lincoln.	22%
	b. The foundation gave Grant's letters praising Lincoln to a museum in Philadelphia.	78%

TABLE 3

Coding schema for Experiments 2 and 3

CODING	EXAMPLE
Prepositional (theme-early)	The teacher read Andy's embarrassing note to Terry to the entire class.
double object (goal-early)	The teacher read the entire class Andy's embarrassing note to Terry.
shifted prepositional (goal-early)	The teacher read to the entire class Andy's embarrassing note to Terry.

TABLE 4

Lexical Bias categorizations of verbs.

Low	High
allot	buy
assign	email
bring	fax
carry	give
mail	hand
radio	read
sell	show
send	tell
ship	

HIGH = frequently occur in the goal-early (double-object) construction.

TABLE 5

Regions analyzed for prosodic breaks in Experiment 3

Region	Example
Verb (immediately after main verb)	<p>AMB: The foundation sent [BREAK] Grant’s letters to Lincoln to a museum.</p> <p>UNAMB: The foundation sent [BREAK] Grant’s letters praising Lincoln to a museum</p>
Embedded phrase (before embedded PP / VP)	<p>AMB: The foundation sent Grant’s letters [BREAK] to Lincoln to a museum.</p> <p>UNAMB: The foundation sent Grant’s letters [BREAK] praising Lincoln to a museum.</p>
Goal PP (immediately before goal PP)	<p>AMB: The foundation sent Grant’s letters to Lincoln [BREAK] to a museum.</p> <p>UNAMB: The foundation sent Grant’s letters praising Lincoln [BREAK] to a museum.</p>

TABLE 6

Proportion of PP-shifted responses in the Goal-early category

	EXPERIMENT 2	EXPERIMENT 3
UNAMBIGUOUS	6%	11%
AMBIGUOUS	17%	22%

TABLE 7

% theme-early (prepositional) constructions used, depending on ambiguity manipulation

	Experiment 2			Experiment 3		
	Unamb.	Amb.	Difference	Unamb.	Amb.	Difference
PP used in the unamb. condition ("Andy's embarrassing note about Terry")	21%	46%	25%	35%	59%	24%
No PP used in unamb. Condition ("Grant's letters praising Lincoln")	11%	16%	4%	22%	41%	19%

Note: Shifted PP responses are not included here (c.f. Pickering, Branigan, & McLean, 2002),

but this does not affect the results. Participants missing data in either condition are excluded from this comparison.

FIGURE CAPTIONS

Figure 1: Experiment 2 Results (Lexical Bias x Ambiguity).

Figure 2: Experiment 2 Results (Weight x Ambiguity).

Figure 3: Experiment 3 Results (Lexical Bias x Ambiguity)

Figure 4: Experiment 3 Results (Weight x Ambiguity)

FIGURE 1

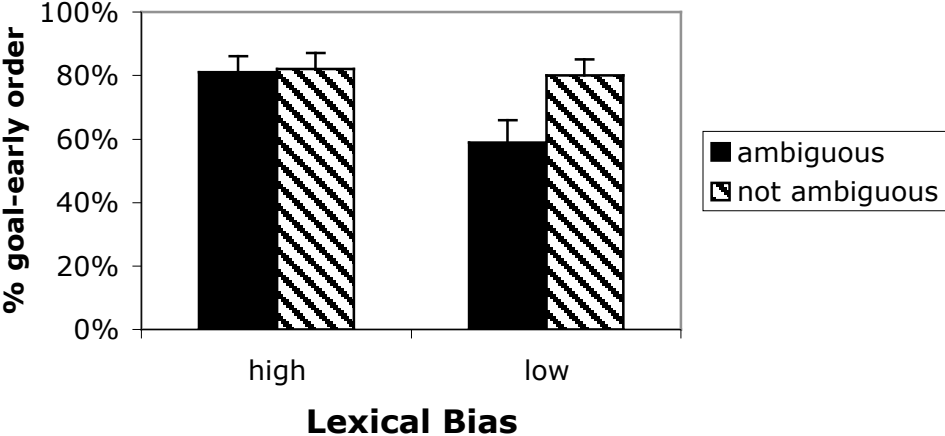


FIGURE 2

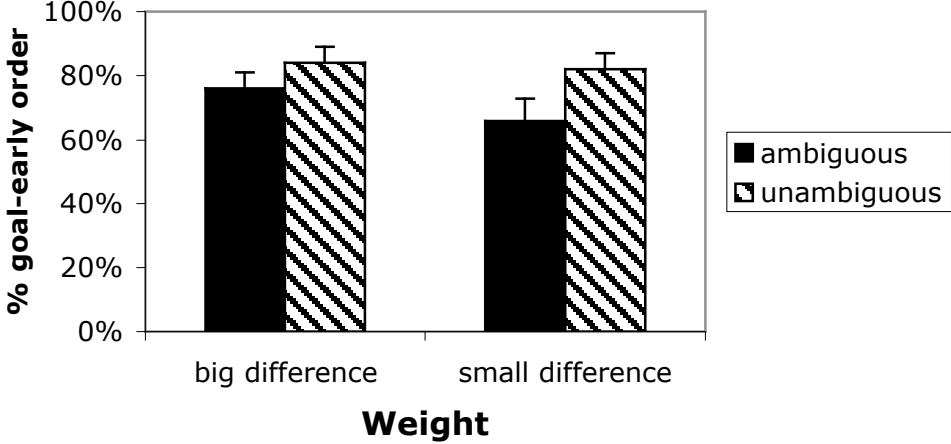


FIGURE 3

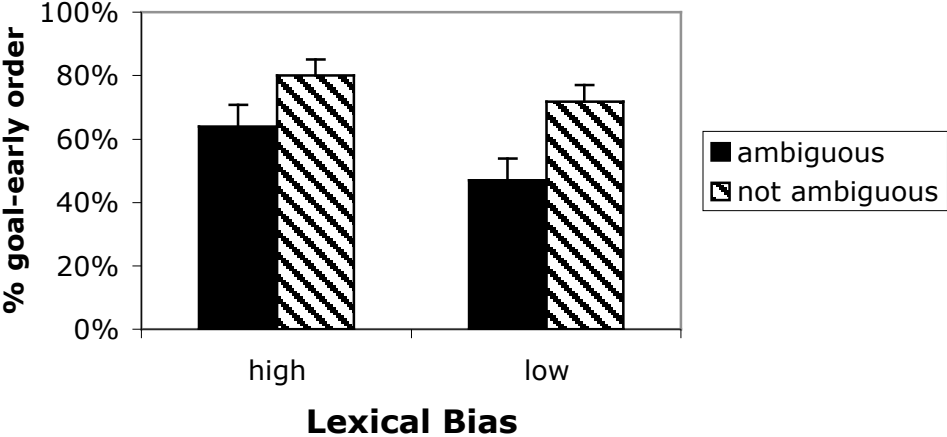


FIGURE 4

