Journal of Memory and Language **40,** 263–279 (1999)
Article ID jmla.1998.2612, available online at http://www.idealibrary.com on IDEAL®

Disambiguation Preferences in Noun Phrase Conjunction Do Not Mirror Corpus Frequency

Edward Gibson

Massachusetts Institute of Technology

and

Carson T. Schütze

University of California, Los Angeles

The results of two self-paced reading studies of a syntactic ambiguity involving conjoined noun phrases to three potential noun phrase sites were compared to the corpus frequencies of the resolutions of the same ambiguity. The reading times for the attachment to the first noun phrase were faster than for the attachment to the second noun phrase, but, to the extent that any differences were observed in the corpus frequencies, attachments to the second noun phrase were more frequent. We therefore argue that the sentence comprehension mechanism is not using corpus frequencies in arriving at its preference in this ambiguity, and hence the decision principles of sentence comprehension and sentence production must be partially distinct. It is proposed that there is a factor operative in sentence comprehension that is not operative in sentence production, and this factor favors attachment to the first noun phrase.

According to an influential proposal by Don Mitchell, Fernando Cuetos, and their colleagues, initial parsing preferences in syntactically ambiguous structures are determined by people's exposure to similar structures in the past (Cuetos & Mitchell, 1988; Mitchell & Cuetos, 1991; Mitchell, 1994; Cuetos, Mitchell, & Corley, 1996; Mitchell, Cuetos, Corley, & Brysbaert, 1996; Mitchell & Brysbaert, 1988; cf. related proposals by Christiansen, 1996; Tabor, Juliano, & Tanenhaus, 1997). Under this proposal—the exposure-based or "tuning" hypothesis—people are assumed to tabulate the resolutions of ambiguities as the ambiguities are encountered, with the result that the most frequently occurring resolution of an ambiguity is the resolution that people prefer. The exposure-based hypothesis was originally put forward in order to account for parsing preference differences between Spanish and English in relative clause (RC) attachment ambiguities such as the following:

(1) a. El periodista entrevisto a [$_{NP_1}$ la hija del [$_{NP_2}$ coronel]] [$_{CP}$ que tuvo el accidente]

b. The journalist interviewed [$_{NP_1}$ the daughter of [$_{NP_2}$ the colonel]] [$_{CP}$ who had had the accident].

In examples like these, Spanish speakers prefer attachment to the first (high) noun phrase (NP) site, whereas English speakers prefer attachment to the second (low) NP site inside the prepositional phrase (PP) (Cuetos & Mitchell, 1988;

We thank the following people for helpful discussions on presentations of this work: the audience at the 9th CUNY Human Sentence Processing Conference, Maria Babyonyshev, Tom Bever, Martin Corley, Danny Fox, Don Mitchell, Edson Miyamoto, Janet Nicol, Neal Pearlmutter, Colin Phillips, Ariel Salomon, Gregg Solomon, Michael Spivey-Knowlton, James Thomas, Susanne Tunstall, Tom Wasow, and two anonymous reviewers. All remaining errors are our own. The second author was supported by the Research Training Grant "Language: Acquisition and Computation" awarded by the National Science Foundation (U.S.) to the Massachusetts Institute of Technology (DIR 9113607) and by a fellowship from the Social Sciences and Humanities Research Council of Canada. Additional funding for the project was provided by the JST/MIT joint international research project "Mind/Articulation."

Address reprint requests to Edward Gibson, E10-034C, MIT, Cambridge, MA 02139. Fax: (617) 253-9767. E-mail: gibson@psyche.mit.edu.

0749-596X/99 $30.00

Clifton, 1988; Mitchell & Cuetos, 1991). (The first attachment site is referred to as the "high" attachment site because it is the site that appears higher in the tree structure of the complex NP "la hija del coronel" / "the daughter of the colonel." For similar reasons, the second site is referred to as the "low" site.) Mitchell, Cuetos and colleagues hypothesized that the reason for the difference between the English and the Spanish preferences is that there is a difference in the relative frequencies of the resolutions of similar ambiguities in the input that English and Spanish speakers are exposed to (see Gilboy, Sopena, Clifton, & Frazier, 1995; Frazier & Clifton, 1996; Gibson, Pearlmutter, Canseco-Gonzalez, & Hickok, 1996; Hemforth, Konieczny, & Scheepers, in press; Sauerland & Gibson, 1998, for alternative explanations of the cross-linguistic attachment preference difference). Cuetos et al. (1996) reported that this is the case in a small-scale study of Spanish and English corpora. In their analyses of instances of two-site RC attachments, Cuetos et al. found that 60% of the RCs in the Spanish examples attached to the high site, whereas only 38% of the RCs in the English examples attached to the high site, as expected under the exposure-based hypothesis.

In contrast to the high-attachment preference found for RC attachments to one of two preceding NP sites in Spanish, there is a low attachment preference when there are three preceding NP sites, as indicated by longer reading times on the region initiated by the disambiguating verb in the RC in (2) (Gibson, Pearlmutter, Canseco-Gonzalez, & Hickok, 1996; Gibson, Pearlmutter, & Torrens, in press):

(2) a. High attachment:
Un alumno insultó a las secretarias del profesor de la clase que no gustaron a los estudiantes.
b. Middle attachment:
Un alumno insultó a la secretaria de los profesores de la clase que no gustaron a los estudiantes.
c. Low attachment:
Un alumno insultó a la secretaria del profesor de las clases que no gustaron a los estudiantes.
"An alumnus insulted the secretary(ies) of the professor(s) of the course(s) that were disliked by the students."

The reading time studies indicate that attachment to the low site in (2c) is the easiest to make, followed by attachment to the high site in (2a), with the middle site (2b) the most difficult of the three. Gibson, Pearlmutter, et al. did not have evidence about the relative frequencies of the disambiguations, so they did not test the exposure-based hypothesis. The present paper presents evidence about the self-paced reading of a closely related ambiguity in English for which there is good evidence about relative frequency: an attachment conjoining an NP with one of three previous NPs, as depicted in Fig. 1.

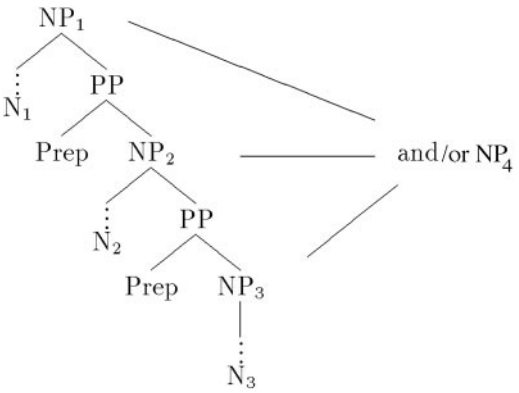

FIG. 1. Ambiguous attachment of an NP to three prospective NP conjunction sites.

Evaluating the exposure-based hypothesis with respect to this kind of ambiguity requires a specific hypothesis about the kinds of frequencies that are tabulated. In their corpus analyses, Cuetos et al. hypothesized that the human sentence processing mechanism tabulates frequencies at the level of a syntactic construction of the form "NP_1 Prep NP_2 RC." However, many other exposure-based "grain"-sizes are logically possible, each of which may make different predictions about parsing preferences. An alternative to the Cuetos et al. syntactic-construction-based frequency proposal is one that is lexically based, with the consequence that frequencies are tabulated at the world level (MacDonald, Pearlmutter, & Seidenberg, 1994; Trueswell, Tanenhaus, & Garnsey, 1994), with no construction-specific frequencies. There is much evidence that people tabulate lexical frequencies in ambiguity resolution (e.g., Tabossi, Colombo, & Job, 1987; MacDonald, 1993, 1994; Trueswell, Tanenhaus, & Kello, 1993; Juliano & Tanenhaus, 1994; MacDonald et al.,

1994; Spivey-Knowlton & Sedivy, 1995; Trueswell, 1996; Garnsey, Pearlmutter, Myers, & Lotocky, 1997; Tabor et al., 1997); however, the lexical exposure-based hypothesis is not sufficient on its own to account for the preferences in the kind of ambiguity being explored here. In particular, the lexical exposure-based hypothesis cannot account for the results of the Gibson et al. (in press) Spanish self-paced reading study. In this experiment, attachment to the low site was preferred in the three-NP-site ambiguity in (2), but attachment to the high site was preferred in a closely related two-NP-site ambiguity, formed by omitting the first of the three prospective NP attachment sites from the three-site examples:

(3) a. High attachment: Un alumno insultó a los profesores de la clase que no gustaron a los estudiantes.
b. Low attachment: Un alumno insultó al profesor de las clases que no gustaron a los estudiantes.
"An alumnus insulted the professor(s) of the course(s) that were disliked by the students."

These results are not compatible with a purely lexically based approach, because both ambiguities involve the same attaching phrase and two of the same potential NP sites immediately preceding it, yet one NP site is favored in the two-site ambiguity (e.g., "profesores" in (3)), whereas the other site is favored in the three-site ambiguity (e.g., "clases" in (2)). In order to account for both the two-site and the three-site data under an exposure-based hypothesis, frequencies must be tabulated at a larger grain than lexical items. Furthermore, tabulating the resolution frequencies of two-NP-site ambiguities alone is not adequate to account for these results, because the attachment preferences vary depending on the number of sites present: the first site is preferred in two-site cases, whereas the last site is preferred in three-site cases. Thus, three-NP-site ambiguity resolution frequencies must also be tabulated to account for the preferences that Gibson, Pearlmutter, et al. (1996) observed.

Similar to the three-NP-site RC attachment ambiguity, the ambiguity to be investigated in this paper involves an attachment to one of three previous NP sites. The arguments given above make a lexically based or a two-NP-site exposure-driven hypothesis unlikely to account for the preferences in this ambiguity. It is necessary to explore the frequencies of three-NP-site ambiguity resolutions to see what the predictions of the exposure-based hypothesis are. Gibson, Schütze, and Salomon (1996) investigated the resolution of this temporary ambiguity with respect to two one-million-word parsed corpora in the University of Pennsylvania Treebank (Marcus, Santorini, & Marcinkiewicz, 1993): the Brown corpus (Kučera & Francis, 1967) and the Wall Street Journal (WSJ) corpus. They found that attachments to the low NP site (NP_3) were the most frequent, followed by attachments to the middle site (NP_2), with attachments to the high site (NP_1) the least frequent of all. Some examples from the corpora are provided in (4)–(6).

(4) Examples of low attached conjoined NPs from the Brown corpus:
a. [$_{NP_1}$ strong opposition by [$_{NP_2}$ the coalition of [[$_{NP_3}$ Southern Democrats] and [$_{NP_4}$ conservative Republicans]]]]
b. [$_{NP_1}$ the running argument about [$_{NP_2}$ the relative merits of [[$_{NP_3}$ Mays] and [$_{NP_4}$ Mickey Mantle]]]]

(5) Examples of middle attached conjoined NPs from the Brown corpus:
a. [$_{NP_1}$ a fine big actor with [[$_{NP_2}$ a great head of [$_{NP_3}$ blond hair]] and [$_{NP_4}$ a good voice]]]
b. [$_{NP_1}$ correct observance of [[$_{NP_2}$ three hundred major rules of [$_{NP_3}$ ritual]] and [$_{NP_4}$ three thousand minor ones]]]

(6) Examples of high attached conjoined NPs from the Brown corpus:
a. [[$_{NP_1}$ a man in [$_{NP_2}$ an occupation of [$_{NP_3}$ high hazard]]] and [$_{NP_4}$ a woman balanced on a knife-edge between death from tuberculosis and recovery]]
b. [[$_{NP_1}$ the question of [$_{NP_2}$ discrimination in [$_{NP_3}$ housing]]] and [$_{NP_4}$ the part each man present played in it]].

This frequency ordering (low, middle, high) was observed for all grain-sizes that Gibson, Schütze, and Salomon considered. The coarsest grain-size that was considered was a three-NP-site ambiguity for all possible attaching categories, pooling conjoining NPs with RCs, PPs, and other attaching categories. At this level of analysis, low attachments were significantly more frequent than middle attachments in both corpora, and middle attachments were significantly more frequent than high attachments in both corpora. Analyzing just the conjoined NPs pro-

vided the same pattern of results, with the exception that the difference between middle and high attachments did not quite reach significance in the WSJ corpus.

Gibson, Schütze, and Salomon analyzed narrower grain-sizes for the comparison between middle and high attachments, considering only attaching phrases that were not lexically biased in one way or another. For example, items in which one of the NP sites was part of an idiomatic expression, such as "in spite of " or "in connection with," do not allow attachment to the NP forming part of the idiom. Similarly, NPs which are parts of proper names, such as "United States of America," are not possible attachment sites. Examples like these were therefore excluded from consideration for the more fine-grained counts. Relatedly, items that include the preposition "between" or the conjunction introducer "both" create a bias to take a following "and" matching at the same level, so examples including these and related words were excluded. The results of these corpus searches revealed that middle-attached examples were still more frequent than high-attached examples, significantly so in the Brown corpus, nonsignificantly so in the WSJ corpus.

Gibson, Schütze and Salomon considered one final grain-size, subdividing the high and middle attachments in terms of the definiteness of the NP sites. In every subcondition except for the uniformly definite case, there were at least as many middle as high attachments. In the uniformly definite case, more high attachments were observed than middle attachments in both corpora, but the numbers were extremely small. Gibson, Schütze and Salomon did not consider any narrower grain-sizes than these because of the huge quantity of input that such grain-sizes would require in order for the parser to set reliable preferences. At the most narrow grain-size that was considered, there were only around 10 instances from a two-million-word corpus.

Given the corpus frequencies observed by Gibson, Schütze, and Salomon, the exposure-based hypothesis predicts that middle attachments should be easier to process than high attachments in conjoined-NP three-site ambiguities. To test this hypothesis, Gibson, Schütze, and Salomon performed an off-line survey rating the comprehensibility of low, middle, and high attachments of the form in (7):

(7) The salesman ignored a customer with a child with a dirty face and
  a. a wet diaper. [low]
  b. one with a wet diaper. [middle]
  c. one with a baby with a wet diaper. [high]

The attachment site for the conjunction "and" was disambiguated to one of the three NP sites in two ways.[1] First, the completions were disambiguated using plausibility information: each completion was plausible under only one of the three prospective attachments. Second, the second conjunct was manipulated so that it was maximally parallel to the NP that it was conjoined with in terms of length and structure. In particular, the conjoined NPs in the complete structures for each of the three versions of (7) contain the same number of PPs (zero, one, or two, for low, middle, and high conjunction, respectively) and are right branching. Assuming that there is a general preference for conjoined constituents to be maximally parallel, the intended attachments will be the preferred ones.

The results of the Gibson, Schütze, and Salomon survey were that low attachments were rated as least complex, followed by high attachments, with middle attachments rated as most complex, the same pattern of results observed for English RC attachments in three-NP-site ambiguities (Gibson, Pearlmutter, et al., 1996). Thus, in contrast to the prediction of the exposure-based hypothesis, high attachments were rated as easier to process than middle attachments, in spite of the fact that middle attachments were more frequent in the corpus.

One possible explanation for the mismatch between the corpus frequencies and the survey results is that the disambiguation in the surveys may not have been representative of the kinds of structures that were enumerated in the corpus

[1] When the conjunction itself is first processed, higher attachment sites are possible, including VP- and S-level conjunction. The former is ruled out by the word following the conjunction. The latter remains possible in principle until the end of the sentence; however, Frazier (1979) has shown that S-level conjunction is dispreferred relative to NP-conjunction, so participants were probably not pursuing this possibility.

TABLE 1

Frequencies in the Brown and WSJ Corpora of NPs Conjoined to the First or Second of Three Preceding NP Attachment Sites, Where the Conjoined NP Contains a Pronoun as Its Head (e.g., "One," "Ones") Or Its Specifier (e.g., "Its")

	Brown corpus		WSJ corpus	
	Middle	High	Middle	High
All matching structures	9	4	7	3
Filtered structures[a]	5	1	1	2

[a] The frequencies of the matching structures without lexical biases in the three preceding NP attachment sites.

counts. In particular, the middle and high attachments are disambiguated using a pronoun "one." It could be that there are more high attachments which are disambiguated using a pronoun than middle attachments which are disambiguated using a pronoun, in contrast to the corpus counts reported by Gibson, Schütze, and Salomon. A further corpus search reveals that this is not a likely explanation of the observed mismatch. Table 1 presents the counts from the WSJ and Brown corpora of NPs conjoined to the first or second of three preceding NP attachment sites, where the conjoined NP contains a pronoun as its head (e.g., "one," "ones") or in a prehead position (e.g., "its," as in "Steele's comment on Swift's change of parties and its effect on their friendship").

Although the number of structures in the corpora consisting of a pronominal element conjoined with one of three previous NP sites is small overall, the frequency breakdown between middle and high attachments follows the general trend observed by Gibson, Schütze, and Salomon: There are more middle attachments than high attachments. In the counts that include all such structures (including some that have lexical biases in the three preceding NP attachment sites) there are more middle attachments than high attachments in both corpora. In the counts in which lexically biased cases are filtered using the same method as Gibson, Schütze, and Salomon, there are more middle attachments than high attachments in the Brown corpus. The pattern is reversed in the WSJ corpus, but there are only three matches altogether in this case.

Because RCs are often introduced by a lexical relative pronoun (e.g., "which," "who"), it is worthwhile to check whether including RCs in the corpus counts has an effect on the pattern of frequencies. However, there were very few such RCs in the high and middle attachment configurations following three NP sites in either corpus. There were none at all in the Brown corpus, and only four in the WSJ corpus, reduced to two after lexical filtering. Inclusion of these frequencies therefore does not change the observed pattern.

Given the observation that there are more middle attachments than high attachments even when the corpus searches are restricted to items containing pronouns in the disambiguating region, the mismatch between the corpus frequencies and the Gibson, Schütze, and Salomon survey results remains unaccounted for. However, there are a number of problems in interpreting the results of their survey. First, because the Gibson, Schütze, and Salomon study involved an off-line rating task, this study might not reflect the initial stages of sentence parsing. It is possible that people's initial parsing preference is for middle attachment over high attachment for conjoined NPs, but that this preference is eventually overridden by a later processing stage in which high attachment is preferred over middle attachment. Second, the items in the Gibson, Schütze, and Salomon study were not controlled for plausibility. The items were disambiguated using plausibility information, but the different disambiguations within an item were not independently tested for their plausibilities to control for potential differences. Thus, the observed complexity advantage for high attachments over middle attachments could have resulted from the high attachment disambiguations being more plausible than the middle attachment disambiguations. Third, in controlling for parallelism between the conjuncts, length was not controlled: The high attachment completions were longer than the middle attachment completions. The experiments reported here address these concerns using self-paced word-by-word reading of more controlled stimulus sentences.

EXPERIMENT 1

Method

Participants. The participants were 32 native English speakers, students and other affiliates of MIT who were paid for their participation.

Materials. The 12 items used in this experiment were similar to the high- and middle-attachment items in the Gibson, Schütze, and Salomon off-line experiment. The form of the items is as shown in (8). A sample item is provided in (9).

(8) Subject-NP Verb NP_1 $Prep_1$ NP_2 $Prep_2$ NP_3 and (the) one . . .
  a. High attachment: $Prep_3$ NP_4 $Prep_4$ NP_5 . . .
  b. Middle attachment: $Prep_4$ NP_5 . . .

(9) The talkshow host told a joke about a man with an umbrella and one . . .
  a. High attachment:
     about a woman with a dog but hardly anybody laughed.
  b. Middle attachment:
     with a dog but hardly anybody laughed.

The attachment site for the conjoined NP "and (the) one" was disambiguated to the high or middle attachment site in three ways. First, the word "one" requires a contrasting modifier, and there was no modifier available on the low attachment site for 10 of the 12 items, ruling out low attachment for these items. For the remaining two items, low attachment was ruled out by plausibility and parallelism, as discussed below. Second, the preposition following the word "one" was the same as the preposition following the high or middle attachment site, thus biasing the attachment toward the matching site because it is more parallel to its conjoining element in terms of lexical content. For example, the preposition "with" following "one" in (9b) helps disambiguate the attachment toward the middle site because the middle attachment site is followed by a PP initiated by "with."

Third, the completions were disambiguated toward the high or middle attachment using plausibility information. For example, the PP "with a dog" can attach to the middle NP position because it is plausible for a man to have both an umbrella and a dog. However, this PP cannot plausibly attach to the high site, because it makes no sense for a joke to have a dog. Similarly, the PP "about a woman" can attach to the high site, because it is plausible for there to be a second joke, a joke about a woman. This PP cannot plausibly attach to the middle site, because it makes no sense for a man to be about a woman.

The plausibilities of the high and middle attachments were matched using an off-line plausibility rating study. In order to preserve meaning and lexical content in the plausibility survey while removing the target temporary ambiguity, the high and middle attachment examples were transformed into descriptions including lists of two elements, as follows:

(10) a. High attachment plausibility:
     The talkshow host told two jokes: one joke about a man with an umbrella; and a second joke about a woman with a dog.
  b. Middle attachment plausibility:
     The talkshow host told a joke about two men: one man with an umbrella; and a second man with a dog.

Two lists of 20 items, consisting of 10 each from the middle and high attachment conditions, were constructed. Twenty filler descriptions were added to each list, giving a total of 40 items in each list. Forty participants who did not take part in either of the self-paced reading experiments rated the descriptions according to their naturalness in the real world on a scale from 1 (natural) to 7 (unnatural). Of these 20 items, 12 were selected for use in the on-line study based on matched plausibility ratings (high attachment 2.22, standard error = 0.11; middle attachment 2.26, standard error = 0.11), in addition to being matched for word length and word frequency in the disambiguating PP region following "one" (mean length of disambiguating PP region for the high attachment condition: 14.0 characters; mean length of disambiguating PP region for the middle attachment condition: 14.7 characters).

These materials address the problems discussed earlier in interpreting the Gibson, Schütze, and Salomon off-line survey. First, the experiment measured on-line reading times, so its results are likely to reflect early stages of processing. Second, the items were pretested for plausibility, so any reading time differences observed are not due to plausibility differences.

Third, the disambiguating region was controlled for length, so reading time differences in this region are not due to length differences.

There were a total of 10 different prepositions used in the items. Of these, three occurred in both the first and the second preposition positions ("about," "by," "with"), three occurred only as the first preposition ("on," "off," "beside"), and four occurred only as the second preposition ("in," "from," "to," "near").

Because corpus analyses demonstrated that the greater frequency of middle versus high attachments was more consistent across indefinite NP-sites, the items were constructed such that most of the NP attachment sites were indefinite. In particular, 8 of the 12 items contained all indefinite NPs initiated by the indefinite article "a" (or "an"), as in (9). In three items, the high and middle NP attachment sites were indefinite, whereas the low attachment site was definite, initiated by the definite article "the." In these 11 items, the conjoining NP was initiated by the indefinite pronoun "one," matching either the high or the middle sites. In the remaining item, all three sites were definite, initiated by the definite article "the," and the conjoining NP was initiated by the definite expression "the one."

In 19 of the 24 completions, the PP region following the word "one" was three words long, consisting of a preposition, an article, and a noun. In the remaining five completions (three middle-attachment items and two high-attachment items) the disambiguating PP region also contained an adjectival modifier, making it four words long, as in (11b):

(11) Today's newspaper has an article about a movie with a French actor and one

a. High attachment:
about a film with a Spanish actress but there is nothing about the new Walt Disney film.

b. middle attachment:
with a Spanish actress but there is nothing about the new Walt Disney film.

See Appendix A for a complete list of the stimuli along with each item's corresponding mean plausibility rating from the pretest.

Each sentence was followed by a yes/no comprehension question. The comprehension questions for 10 of the 12 experimental items were designed so that it could be determined whether the reader obtained the appropriate interpretation for the sentence. To achieve this goal, the number of objects named by NP_1 was questioned. If high attachment was required, then there were two objects corresponding to NP_1, and one corresponding to NP_2. On the other hand, if middle attachment was required, then there was only one object corresponding to NP_1 and two corresponding to NP_2. For example, in the high attachment completion in (9a), the talkshow host told two jokes (NP_1), and the first joke was about exactly one man (NP_2). In the middle attachment completion in (9b), the talkshow host told only one joke (NP_1), and that joke was about two men (NP_2). The comprehension questions therefore asked whether there were one or two of the objects named by the high attachment site. For example, the question for (9) is (12):

(12) Did the talkshow host tell two jokes?

For the 10 items having this form of comprehension question, 5 asked whether there were two objects indicated by NP_1 (the correct answer is "yes" for high attachment completions and "no" for middle attachment completions), and 5 asked whether there was one object indicated by NP_1 (the correct answer is "no" for high attachment completions and "yes" for middle attachment completions).

Each participant read one version of each test sentence. The 12 items were interspersed with 60 filler sentences of various types, including items from other experiments with unrelated hypotheses. Each participant encountered the sentences in a different pseudorandom order.

Procedure. Participants were timed in a word-by-word self-paced noncumulative moving-window reading task (Just, Carpenter, & Woolley, 1982) controlled by an IBM PS2 computer running Micro-Experimental Laboratory (MEL) software. Participants pressed the space bar to reveal each subsequent word and cause all other words to revert to dashes. At the end of each sentence, a yes/no question appeared on the screen, which participants answered by pressing one of two keyboard keys. Participants were informed by a screen message when they

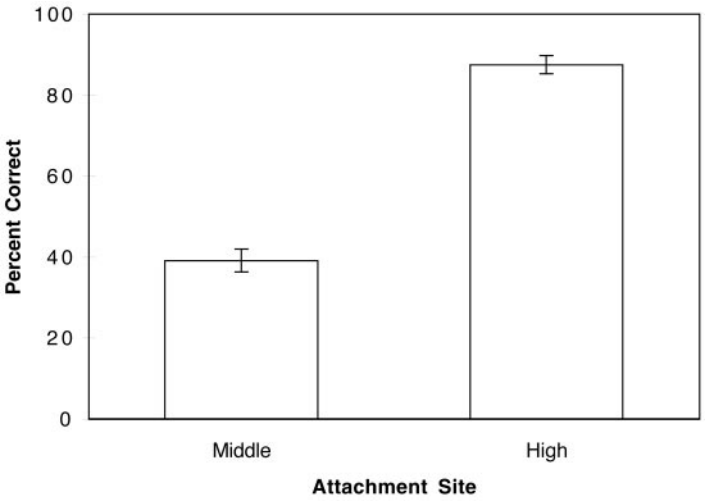

FIG. 2. Comprehension question response accuracy in Experiment 1. The error bars represent standard errors.

answered incorrectly, in order to encourage them to keep paying attention to the content of the sentences.

The experimental trials were preceded by two screens of instructions and eight practice trials. All sentences spanned at least two lines on the screen, and no sentence spanned more than three. Because the MEL software allowed at most 80 characters per line, and because the target items were mostly much longer than this, the disambiguating region occurred on the second line of each target item. To control for potential slowdowns caused by initiating a new line within a sentence, the target items were presented so that a new line was always initiated immediately before the conjoined NP starting with the words "and (the) one . . .," as shown below in (13). The experiment took participants approximately 20 min.

Analysis. We analyzed the comprehension question response accuracy and the reading times. For the purposes of analysis and presentation of the data only, items were separated into five regions as illustrated in (13). The critical region is the disambiguating prepositional phrase following the word "one." As discussed above, this region consisted of three words for 19 of the 24 completions, and four words for the remaining 5 because of the presence of an adjectival modifier.

(13) The talkshow host told a joke / about a man with an umbrella /

a. High attachment:
and one / about a woman / with a dog but hardly anybody laughed.

b. Middle attachment:
and one / with a dog / but hardly anybody laughed.

Results

Comprehension question response accuracy. The response accuracies for the high- and middle-attachment conditions, expressed as percentages, are presented in Fig. 2. Participants were correct in answering questions to the high attachment condition significantly more often than in answering questions to the middle attachment condition (mean for high attachment = 88%; mean for middle attachment = 39%; $F_1(1,31) = 187.4$, $MS_e = 2.00$, $p < .001$; $F_2(1,11) = 24.3$, $MS_e = 5.79$, $p < .001$). The comprehension questions for two middle-attachment condition items were answered extremely poorly, with two or fewer correct answers each across all 32 participants. The extreme difficulty with these questions suggests that the corresponding stimuli were interpreted differently than was intended, so analyses were also performed without these items. Even without these two items, a highly significant difference between high- and middle-attachment remains with respect to comprehension question

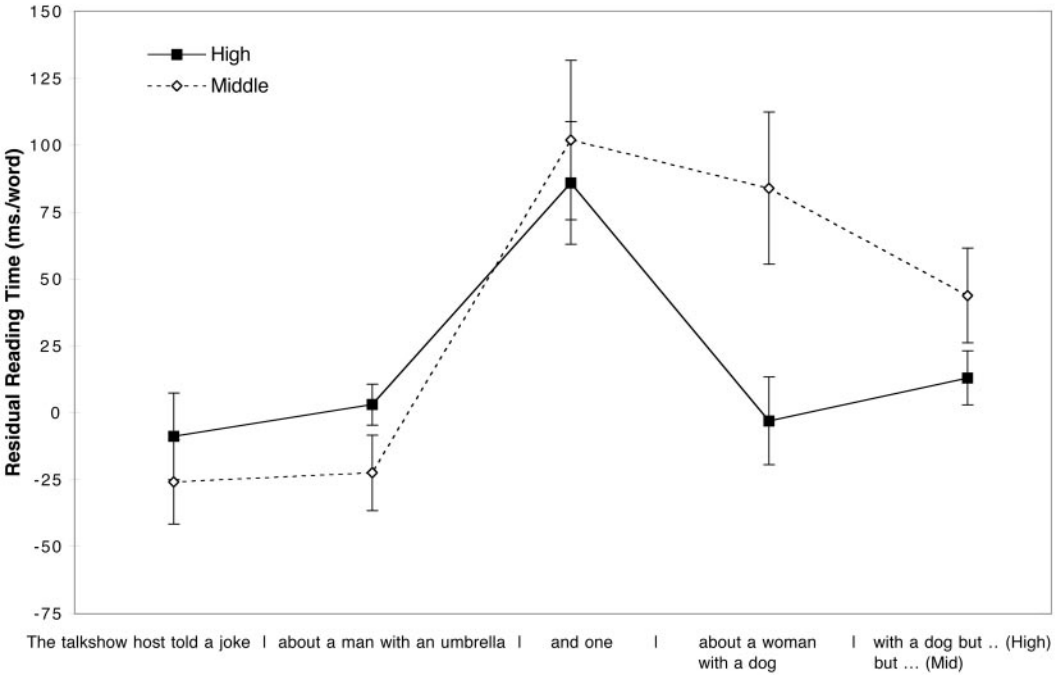

FIG. 3. Residual reading times in Experiment 1. The error bars represent standard errors.

response accuracy (mean for high attachment = 86%; mean for middle attachment = 46%; $F_1(1,31) = 81.9$, $MS_e = 3.03$, $p < .001$; $F_2(1,9) = 17.9$, $MS_e = 4.34$, $p < .005$).

Reading times. We omitted the data from the two items whose comprehension questions received two or fewer correct answers each for the whole experiment. We also omitted all data from participants who answered at most one of the middle-attachment questions correctly, resulting in the omission of the data for 5 of the 32 participants. (No participants had comparable difficulty in answering the high-attachment questions: All participants answered at least three of the six high-attachment questions correctly). For the remaining data, trials on which the question was answered incorrectly were also excluded from the analysis. This removed 30.4% of the remaining data (a 13.3% error rate for high attachments; a 48.4% error rate for middle attachments). In addition to analyzing raw reading times per word, we also analyzed residual reading times per word (Ferreira & Clifton, 1986), derived by subtracting from raw reading times each participant's predicted time to read words of the same length, calculated by a linear regression equation across all sentences in the experiment. Residual reading times and standard errors are displayed in Fig. 3. Raw reading times are presented in Table 2, and Appendix B presents raw reading times for data from all participants and all items, whether or not the comprehension question was answered correctly. The patterns for all of the sets of data are similar.

There were no significant reading time differences between the conditions before the disambiguating region (all ps ≥ 0.15). (Note that although the standard error bars do not overlap in Region 2 in Fig. 3, the reading times are not significantly different in this

TABLE 2

Mean Raw Reading Times per Word (in Milliseconds) for Experiment 1

	Sentence region				
Condition	1	2	3	4	5
Middle	366	349	459	456	428
High	378	376	443	364	393

region ($p = .15$)). On the disambiguating PP region, the high attachment condition was read faster than the middle attachment condition ($F_1(1,26) = 10.29$, $MS_e = 9917$, $p < .005$, $F_2(1,9) = 6.48$, $MS_e = 10878$, $p < .05$). This difference was also significant in this region for the comparison of raw reading times, although only marginally so in the items analysis ($F_1(1,26) = 10.92$, $MS_e = 10336$, $p < .005$; $F_2(1,9) = 4.62$, $MS_e = 17454$, $p = .06$). There was a tendency toward a difference in reading times in Region 5, the region following the disambiguating region. Reading times were slower in this region for the middle attachment condition than for the high attachment condition, but only marginally in the participants analysis ($F_1(1,26) = 3.37$, $MS_e = 3790$, $p = .08$) and nonsignificantly in the items analysis ($F_2(1,9) = 1.78$, $MS_e = 2356$, $p = .21$). The comparison of raw reading times in this region resulted in a significant difference in the participants analysis ($F_1(1,26) = 4.67$, $MS_e = 3697$, $p < .05$), but not in the items analysis ($F_2(1,9) = 2.73$, $MS_e = 3095$, $p = .13$).

In order to evaluate the possibility that plausibility differences among the items might be contributing to the observed differences in the disambiguating region, we tested to see if there was a correlation between the plausibility difference scores obtained in the pretest and the reading time differences on an item-by-item basis, averaging over participants. The resulting correlation was not significant ($r = .18$; $p > .6$), suggesting that reading time differences were not due to plausibility differences among the items.

Discussion

The results of Experiment 1 demonstrate the difficulty that people have in comprehending the middle attachment completions as compared with the high attachment completions, thus replicating the Gibson, Schütze, and Salomon findings. Many of the participants in Experiment 1 could not understand the middle attachment completions well enough to answer the comprehension questions at chance or better. The same group of participants had no comparable difficulty with the high attachment completions. Furthermore, the participants read the middle attachment completions significantly more slowly than the high attachment completions.

Although the results of Experiment 1 are suggestive, a problem remains in interpreting the results. It could be that the difference between high and middle attachments stems partially from the line break which occurred immediately before the conjunction. The presence of the line break led to increased reading times, as shown in Fig. 3. It is possible that the presence of the line break caused participants to close off the lower two attachment sites within the complex NP, with the result that it was possible to conjoin only the whole NP once the new line had been initiated. Experiment 2 addresses this concern.

EXPERIMENT 2

Method

Participants. The participants were 47 native English speakers, students and other affiliates of MIT who were paid for their participation, none of whom participated in Experiment 1 or the plausibility survey described there.

Materials. The target items were the same as in Experiment 1, with one exception: the one item whose prospective attachment sites were all definite. In order to obtain greater uniformity among the items, this item was replaced by an item with uniformly indefinite attachment sites taken from the plausibility pretest described in Experiment 1. The new set of 12 items were still matched for plausibility ratings (high attachment 2.28, standard error = 0.10; middle attachment 2.23, standard error = 0.10), in addition to being matched for word-length and word-frequency in the disambiguating PP region following "one" (mean length of disambiguating PP region for the high attachment condition 13.7 characters; mean length of disambiguating PP region for the middle attachment condition, 14.6 characters).

The comprehension questions for the items were rewritten so that all 12 items had questions whose responses could distinguish whether the reader obtained the appropriate interpretation for the sentence, by asking about the number of

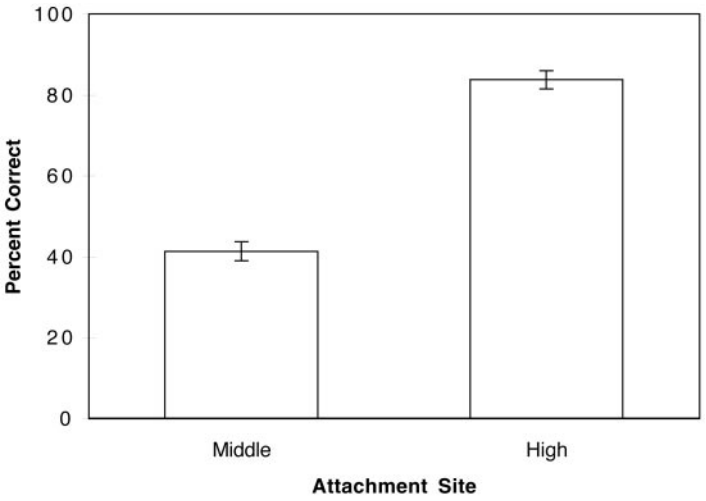

FIG. 4. Comprehension question response accuracy in Experiment 2. The error bars represent standard errors.

objects in the sentence named by the high or middle attachment sites. Six of the items had questions which asked about the number of objects indicated by NP_1, and the other six items had questions which asked about the number of objects indicated by NP_2. Half of each of these sets of six were constructed so that the correct answer for the question was "yes" in the high attachment and "no" in the middle attachment, and the other half were constructed in the opposite manner.

Each participant read one version of each test sentence. The 12 items were interspersed with 68 filler sentences of various types. These included items from other experiments with unrelated hypotheses. Each participant encountered the sentences in a different pseudorandom order.

Procedure. The experimental procedure was the same as that for Experiment 1, except that it was run on a Macintosh Centris computer using custom software which allowed 100 characters per line. This additional screen width made it possible to present regions 1 through 4, including the conjunction region (region 3) and the disambiguating region (region 4) all on the same line. The items all continued on a second line, but only after at least four words of disambiguation were presented on the first line. The experiment took participants approximately 20 min.

Results

One participant was omitted from the analysis because he discovered what the target ambiguity was that was being tested during the experiment. However, his data was similar to that of the other participants in the experiment; including his data has no effect on the results reported here.

Comprehension question response accuracy. The response accuracies for the high- and middle-attachment conditions, expressed as percentages, are presented in Fig. 4. Participants were correct in answering questions to the high attachment condition significantly more often than in answering questions to the middle attachment condition ($F_1(1,45) = 187.5$, $MS_e = 2.20$, $p < .001$; $F_2(1,11) = 48.6$, $MS_e = 2.22$, $p < .001$).

Reading times. As in Experiment 1, we omitted data from participants who answered at most one of the middle-attachment questions correctly, resulting in the omission of the data for 8 of the 46 participants. For the remaining data, trials on which the question was answered incorrectly were also excluded from the analysis. This removed 34.4% of the remaining data (a 15.4% error rate for high attachments, a 53.5% error rate for middle attachments). Residual reading times and standard errors are displayed in Fig. 5. Raw reading times are presented in

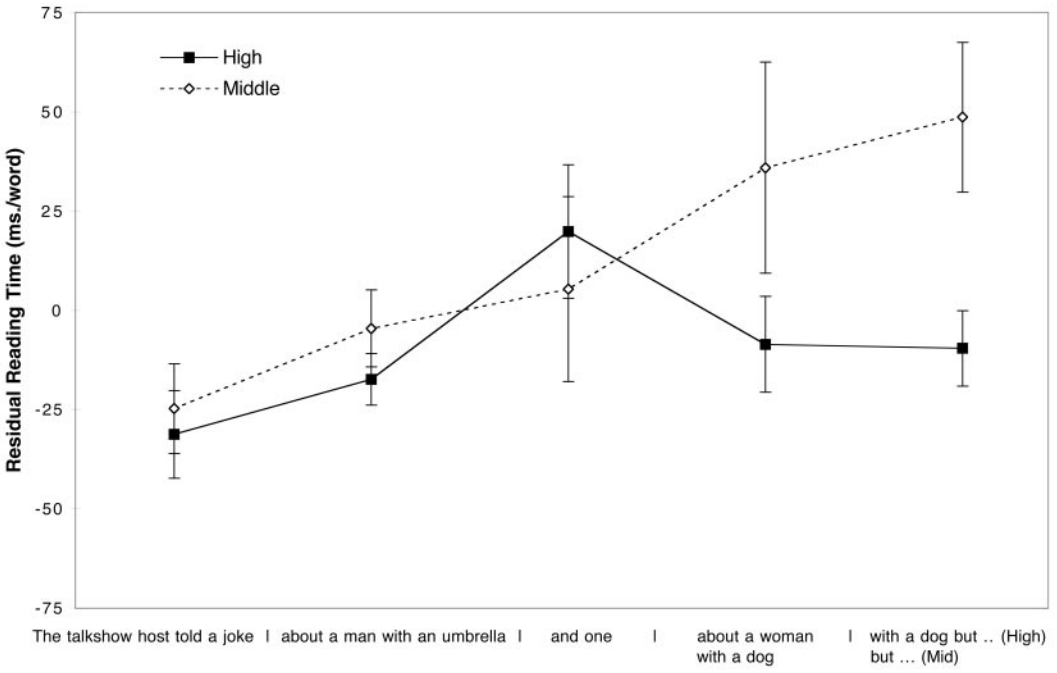

FIG. 5. Residual reading times in Experiment 2. The error bars represent standard errors.

Table 3, and Appendix C presents raw reading times for data from all participants and all items, whether or not the comprehension question was answered correctly.

There were no significant reading time differences between the conditions before the disambiguating region (all ps > 0.3). In the disambiguating region, the high attachment condition was read faster than the middle attachment condition, significantly by participants ($F_1(1,37) = 4.37$, $MS_e = 8618$, $p < .05$), marginally by items ($F_2(1,11) = 3.87$, $MS_e = 5907$, $p = .08$). This difference was also significant in this region for the comparison of raw reading times ($F_1(1,37) = 4.42$, $MS_e = 8754$, $p < .05$; $F_2(1,11) = 7.34$, $MS_e = 4304$, $p < .05$). The high attachment completions were also read significantly faster in the region containing the rest of the sentence, including the line break ($F_1(1,37) = 9.55$, $MS_e = 6760$, $p < .005$; $F_2(1,11) = 10.81$, $MS_e = 3130$, $p < .01$). The comparison of raw reading times in this region also resulted in a significant difference ($F_1(1,37) = 12.70$, $MS_e = 6696$, $p < .001$; $F_2(1,11) = 18.76$, $MS_e = 2737$, $p < .001$).

TABLE 3

Mean Raw Reading Times per Word (in Milliseconds) for Experiment 2

	Sentence region				
Condition	1	2	3	4	5
Middle	391	416	414	489	513
High	381	398	424	413	420

As in Experiment 1, we also tested to see if there was a correlation between the plausibility difference scores obtained in the pretest for Experiment 1 and the reading time differences on an item-by-item basis, averaging over participants. Similar to the results from Experiment 1, the resulting correlation was not significant ($r = .05$; $p > .8$), suggesting that reading time differences were not due to plausibility differences among the items.

Discussion

The same pattern of data was observed in Experiment 2 and in Experiment 1. The middle

attachment completions were much harder to understand than the high attachment completions, as indicated by the question–answering data. In addition, participants read the disambiguating region more slowly for the middle attachment completions than for the high attachment completions. Furthermore, the reading time differences observed here are not related to processing sentences across line breaks, because the line break occurred after the disambiguating region for the items in this experiment. It is concluded that the middle attachment completions are harder to understand than the high attachment completions.

GENERAL DISCUSSION

The results of Experiments 1 and 2 demonstrate that middle conjunction attachments are more difficult to process than high conjunction attachments. In contrast, middle attachments are more frequent than high attachments in the corpora that were analyzed by Gibson, Schütze, and Salomon. This pattern of results is not expected under the Mitchell and Cuetos exposure-based framework for any grain-size that was explored, nor is it predicted by other exposure-based accounts of ambiguity resolution in sentence comprehension, including lexically based constraint-satisfaction proposals (e.g., MacDonald et al., 1994; Trueswell et al., 1994) or connectionist-network accounts (Christiansen, 1996; Tabor et al., 1997). Of course, the fact that the grain sizes investigated by Gibson, Schütze, and Salomon fail to account for the complexity preferences is not a proof that there is no grain size that makes the right predictions. In principle, there could be other grain sizes that have not yet been identified which allow exposure-based explanations of the reading-time findings. Another possible explanation for the noncorrelation between frequency and complexity is that the input from which the behavioral patterns are learned may be different from the corpora that were analyzed. That is, it could be that the parsed corpora that were analyzed are not representative of typical English texts. Some doubt is cast on this possibility by the observation that the frequency distributions for the constructions in question are quite similar in the two corpora. A related potential explanation of the noncorrelation is that the exposure-based preferences for a given language might be based on spoken language rather than written language, so that the analysis of written texts may not be directly relevant to the issue in question. This possibility should be taken seriously, especially if parsing preferences are established while the grammar of a language is being learned, before children are reading very much. However, until large corpora of adult-to-child speech are available, there is no way to assess whether they differ from the written corpora in terms of the relative frequencies under discussion.

In any case, the discrepancy between comprehension complexity and corpus frequency for the corpora under consideration requires an explanation. While exposure may play a considerable role in sentence comprehension, other factors also seem to apply in the constructions under consideration. We propose that, because of working memory constraints, both sentence production and comprehension are sensitive to locality considerations, such that more local relationships are easier to process than longer distance ones (Gibson, 1998; cf. Ross, 1967; Bever, 1970; Kimball, 1973; Frazier & Fodor, 1978; Frazier, 1979; Hawkins, 1990, 1994; Gibson, 1991; Stevenson, 1994). There is much evidence for a locality principle in sentence comprehension (see Gibson, 1998, for a summary). Under locality considerations alone, low attachment structures should be favored over middle attachment structures in sentence comprehension, and middle attachment structures should be favored over high attachment structures, because the attachment site is most local in low attachments, and least local in high attachments.

Corpus analyses suggest that a locality principle is also active in sentence production: the most complex components of a sentence generally occur sentence-finally in head-first languages, thereby maximizing local relationships. Hawkins (1994) provides data from English and other head-initial languages showing that longer ("heavier") items do generally occur later when there is a choice of word orders. Gibson et al. (1996) provide corpus analyses of English conjoined NPs in the Penn Treebank, and show that

the second conjunct is significantly longer than the first on average. Furthermore, Gibson, Schütze, and Salomon show that the greater the length discrepancy between two conjuncts, the greater the likelihood that the longer one will occur second, as expected under the locality hypothesis.

We suggest that there is more effort involved in producing structures with nonlocal attachments than with local attachments, other factors being equal, because the earlier attachment sites decay in working memory as more material is produced. A low attachment structure should therefore be the easiest to produce among the three possibilities for the three-NP-site ambiguity under consideration here, followed by the middle attachment structure, with the high attachment structure being hardest. Assuming that structures that are harder to produce are produced less frequently, locality in production predicts the observed frequencies in the corpora for this ambiquity (cf. Gibson & Pearlmutter, 1994; Frazier, 1995; Stevenson & Merlo, 1997). Locality is of course but one factor applying in production, in addition to lexical constraints and other factors such as parallelism, favoring balanced attachments. (see Gibson, Schütze, and Salomon for an alternative explanation for the corpus frequencies).

In order to account for the discrepancy between this pattern of production and the comprehension complexity results, we suggest that an additional factor is involved in sentence comprehension, one which is not applicable in sentence production. This factor favors high attachments over all others in cases of ambiguity such as the constructions under consideration. One proposal for such a factor is *predicate proximity* (Gibson et al., 1996), which states that attachments should be made as close as possible to the head of a predicate phrase (typically a verb phrase). The motivation given for this proposal was that because all grammatical utterances have a predicate (verb) at their core, a verb is kept in memory more strongly than other categories. Because there are cross-linguistic differences in attachment preferences involving RCs and preceding NP sites, Gibson, Pearlmutter, et al. hypothesized that the strength of predicate proximity is determined by exposure to the input language, as determined by the average distance from the head of a predicate (verb) to its arguments (e.g., subject and object). It was proposed that languages with larger average distances between verbs and their arguments will require that predicates be more highly activated to permit the necessary attachments over longer distances. This larger predicate activation then results in a greater preference to attach close to the predicate in an ambiguity and thus to a larger influence of predicate proximity in such languages.

An alternative candidate for this factor is the Hemforth et al. (in press) *anaphoric binding* hypothesis, which specifies that the parser initiates a search for the appropriate referent for a pronoun when the pronoun is first encountered. This process influences RC attachment because a search is initiated for the referent of a relative pronoun heading an RC (e.g., "who" or "which"), just as for any other pronoun. This process would also influence the conjoined NP attachments in the kinds of items in the Gibson, Schütze, and Salomon acceptability rating experiment and the two self-paced reading experiments presented here, because the conjoined NP in the middle and high attachment items was initiated with a pronominal element "(the) one." Hemforth et al. argue that the parser prefers to coindex pronouns with elements which are part of the main assertion of a sentence (cf. De Vincenzi & Job, 1995; Frazier, 1990; Frazier & Clifton, 1996; Gibson, Pearlmutter et al. 1996), and thus coindexations with the structurally highest NPs are preferred. The combined effects of locality and a high-favoring factor along either of these lines will make the middle attachment site least preferred in comprehension, because it is favored by neither principle. We suggest that the high favoring factor is a disambiguation mechanism, and hence does not apply in production, where the intended meaning is known.

CONCLUSIONS

This paper has evaluated the exposure-based framework with respect to conjunctions of noun phrases in constructions with three available NP sites in English. Although the experiments demonstrated a preference for high-site attachments

over middle-site attachments, no frequencies in the corpora reflect this complexity ordering for any of the grain-sizes that were evaluated. A purely exposure-based framework is difficult to reconcile with these data. The results can be accounted for by a combination of two factors (in addition to the possibility of exposure-based influences): (1) a locality principle for both production and comprehension and (2) a factor favoring high attachments—such as predicate proximity or anaphoric binding—in ambiguity resolution in sentence comprehension only.

APPENDIX A

Items for Experiments 1 and 2[2]

The high attachment versions of the items below include the prepositional phrase in parentheses. The middle attachment versions omit the parenthetical prepositional phrase. The plausibility rating for the disambiguated form of each item is presented in parentheses following each item. These ratings were obtained in the pretest described in Experiment 1. The plausibility scale goes from 1 (natural) to 7 (unnatural).

1. The kids' magazine printed a story about a haunted house near a pond and one (about an old mansion) near a river because Halloween was coming soon. (High, 2.25; Middle, 2.05)
2. The detective memorized the address on the note to the victim and the one (on the letter) to the suspect before going to look for witnesses. (High, 2.45; Middle, 2.70)
3. At the exhibit, Mary's aunt admired a photo of a child on the beach and one (of a man) in the ocean but she decided not to buy anything. (High, 2.05; Middle, 2.55)
4. The sportswriter wrote a column about a soccer team from the suburbs and one (about a baseball team) from the city for the paper's Sunday magazine. (High, 2.00; Midle, 2.55)
5. The talkshow host told a joke about a man with an umbrella and one (about a woman) with a dog but hardly anybody laughed. (High, 2.80; Middle, 2.45)
6. The journalist wrote a report on a lecture by a prominent businessman and one (on a speech) by a famous economist for a political newsletter. (High, 2.15; Middle, 1.75)
7. The careful student noticed a notepad beside a book about the Revolution and one (beside a text) about the Civil War when she was doing research. (High, 2.25; Middle, 2.45)
8. The review panel heard a presentation by a scientist from the midwest and one (by a professor) from the south before deciding which proposals to fund. (High, 1.85; Middle, 1.75)
9. The costume designer drew a sketch of a dress with a zipper and one (of a coat) with a belt to show the director. (High, 2.50; Middle, 2.50)
10. Today's newspaper has an article about a movie with a French actor and one (about a film) with a Spanish actress but there is nothing about the new Walt Disney film. (High, 1.90; Middle, 2.05)
11. The salesman ignored a customer with a baby in a stroller and one (with an infant) in a carriage while he helped a rich-looking woman. (High, 2.65; Middle, 2.70)
12. The supervisor was given a note about a meeting with a supplier and one (about an appointment) with a customer when he came in on Tuesday morning. (High, 1.85; Middle, 1.85)

[2] The items for Experiment 2 were identical to those in Experiment 1, with the exception that item 2 was replaced with the following item: The travel agency organized a trip to a village in the mountains and one (to a town) in the plains but very few people signed up (High, 3.05; Middle, 2.10).

APPENDIX B

Mean Raw Reading Times per Word (in Milliseconds) for Experiment 1 for Data from All Subjects and All Items

	Sentence region				
Condition	1	2	3	4	5
Middle	364	352	410	387	385
High	381	367	433	357	358

APPENDIX C

Mean Raw Reading Times per Word (in Milliseconds) for Experiment 2 for Data from All Subjects and All Items

	Sentence region				
Condition	1	2	3	4	5
Middle	365	364	377	390	388
High	368	360	379	364	362

REFERENCES

Bever, T. G. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 279–362). New York: Wiley.

Christiansen, M. H. (1996). *Recursive sentence structure in connectionist networks.* Paper presented at the ninth annual CUNY Conference on Human Sentence Processing, CUNY, March, 1996.

Clifton, C., Jr. (1988). *Restriction on late closure: Appearance and reality.* Paper presented at the 6th Australian

Language and Speech Conference, University of New South Wales, August 19–21.

Cuetos, F., & Mitchell, D. C. (1988). Cross-linguistic differences in parsing: Restrictions on the use of the Late Closure strategy in Spanish. *Cognition,* **30,** 73–105.

Cuetos, F., Mitchell, D. C., & Corley, M. M. B. (1996). Parsing in different languages. In M. Carreiras, J. E., García-Albea, & N. Sebastián-Gallés (Eds.), *Language processing in Spanish* (pp. 145–187). Mahwah, NJ: Erlbaum.

De Vincenzi, M., & Job, R. (1995). An investigation of Late Closure: The role of syntax, thematic structure, and pragmatics in initial and final interpretation. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **21,** 1305–1321.

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

Frazier, L. (1979). *On comprehending sentences: Syntactic parsing strategies.* Univ. of Connecticut doctoral dissertation. Bloomington, IN: Indiana University Linguistics Club.

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. (1995). Constraint satisfaction as a theory of sentence parsing. *Journal of Psycholinguistic Research,* **24,** 437–468.

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

Frazier, L., & Fodor, J. D. (1978). The Sausage Machine: A new two-stage parsing model. *Cognition,* **6,** 291–325.

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

Gibson, E. (1991). *A computational theory of human linguistic processing: Memory limitations and processing breakdown.* Unpublished doctoral dissertation, Carnegie Mellon University, Pittsburgh, PA.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition,* **68,** 1–76.

Gibson, E., & Pearlmutter, N. J. (1994). A corpus-based analysis of psycholinguistic constraints on prepositional-phrase attachment. In C. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing* (pp. 181–198). Hillsdale, NJ: Erlbaum.

Gibson, E., Pearlmutter, N., Canseco-Gonzalez, E., & Hickok, G. (1996). Recency preference in the human sentence processing mechanism. *Cognition,* **59,** 23–59.

Gibson, E., Pearlmutter, N., & Torrens, V. (in press). Recency and lexical preferences in Spanish. *Memory and Cognition.*

Gibson, E., Schütze, C. T., & Salomon, A. (1996). The relationship between the frequency and the complexity of linguistic structure. *Journal of Psycholinguistic Research,* **25,** 59–92.

Gilboy, E., Sopena, J. M., Clifton, C., Jr., & Frazier, L. (1995). Argument structure and association preferences in Spanish and English complex NPs. *Cognition,* **54,** 131–167.

Hawkins, J. A. (1990). A parsing theory of word order universals. *Linguistic Inquiry,* **21,** 223–262.

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

Hemforth, B., Konieczny, L., & Scheepers, C. (in press). Syntactic attachment and anaphor resolution: Two sides of relative clause attachment. In M. Crocker, M. Pickering, & C. Clifton, Jr. (Eds.), *Architectures and mechanisms for language processing.* Cambridge, UK: Cambridge Univ. Press.

Juliano, C., & Tanenhaus, M. K. (1994). A constraint-based lexical account of the subject/object attachment preference. *Journal of Psycholinguistic Research,* **23,** 459–472.

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General,* **3,** 228–238.

Kimball, J. (1973). Seven principles of surface structure parsing in natural language. *Cognition,* **2,** 15–47.

Kučera, H., & Francis, W. N. (1967). *Computational analysis of present day American English,* Providence, RI: Brown Univ. Press.

MacDonald, M. C. (1993). The interaction of lexical and syntactic ambiguity. *Journal of Memory and Language,* **32,** 692–715.

MacDonald, M. C. (1994). Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes,* **9,** 157–202.

MacDonald, M. C. (1997). Lexical representations and sentence processing: An introduction. *Language and Cognitive Processes,* **12,** 121–136.

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

Marcus, M. P., Santorini, B., & Marcinkiewicz, M. A. (1993). Building a large annotated corpus of English: The Penn Treebank. *Computational Linguistics,* **19,** 313–330.

Mitchell, D. C. (1994). Sentence parsing. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 375–409). San Diego, CA: Academic Press.

Mitchell, D. C., & Brysbaert, M. (1998). Challenges to recent theories of crosslinguistic variation in parsing: Evidence from Dutch. In D. Hillert (Ed.), *Syntax and semantics: A crosslinguistic perspective* (pp. 313–335). San Diego, CA: Academic Press

Mitchell, D. C., & Cuetos, F. (1991). The origins of parsing strategies. *Conference proceedings: Current issues in natural language processing* (pp. 1–12). Univ. of Texas at Austin, Austin, TX.

Mitchell, D. C., Cuetos, F., Corley, M. M. B., & Brysbaert, M. (1996). Exposure-based models of human parsing:

Evidence for the use of coarse-grained (non-lexical) statistical records. *Journal of Psycholinguistic Research*, **24,** 469–488.

Ross, J. R. (1967). *Constraints on variables in syntax*. Doctoral dissertation, MIT, Cambridge, MA.

Sauerland, U., & Gibson, E. (1998). *Case matching in relative clause attachment*. Paper presented at the eleventh annual CUNY sentence processing conference, Rutgers University, New Brunswick, NJ, March, 1998.

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

Stevenson, S. (1994). Competition and recency in a hybrid network model of syntactic disambiguation. *Journal of Psycholinguistic Research*, **23,** 295–322.

Stevenson, S., & Merlo, P. (1997). Lexical structure and parsing complexity. *Language and Cognitive Processes*, **12,** 349–399.

Tabor, W., Juliano, C., & Tanenhaus, M. K. (1997). Parsing in a dynamical system: An attractor-based account of the interaction of lexical and structural constraints in sentence processing. *Language and Cognitive Processes*, **12,** 211–272.

Tabossi, P., Colombo, L., & Job, R. (1987). Accessing lexical ambiguity: Effects of context and dominance. *Journal of Psychological Research*, **49,** 161–167.

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

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

Trueswell, J. C., Tanenhaus, M. K., & Kello, C. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **19,** 528–553.

(Received January 16, 1998)
(Revision received July 31, 1998)