On the Strength of the Local Attachment Preference

Colin Phillips[1,3] and Edward Gibson[2]

This paper investigates the strength of the local attachment preference in syntactic ambiguity resolution, based on a study of a novel ambiguity for which the predictions of local attachment contrast with the predictions of a wide range of other ambiguity resolution principles. In sentences of the form "Because Rose praised the recipe I made ..." we show that the ambiguous clause "I made" is preferentially attached as a relative clause under some circumstances, as predicted by local attachment, and preferentially attached as a matrix clause under other circumstances. The implications for accounts of locality in parsing are discussed.

THE LOCALITY PUZZLE

This paper is a progress report on our work which investigates the strength and the generality of the *local attachment preference* (see Gibson, Pearlmutter, Canseco-Gonzalez, & Hickok 1996; Phillips 1995, 1996). We use this term to refer in a theory neutral way to whatever underlies the inter-

pretive preference that speakers of English share when presented with a
sentence like (1a). The first reading that becomes available is one in which
the locative adverbial PP *in the garden* associates with the lower clause,
such that it is the frog's singing that is happening in the garden, and not the
announcement that Alice made about the frog. In other words, the adverbial
is preferentially attached to local material in the tree rather than to nonlocal
material. The local attachment preference is a pervasive phenomenon in
parsing, and some more well-known cases of local attachment from Kimball
(1973) are shown in Sentences (1c) to (1e) (the # mark indicates that a
sentence is typically misanalyzed or hard to understand).

(1) a. Alice said the frog was singing in the garden.

    b.
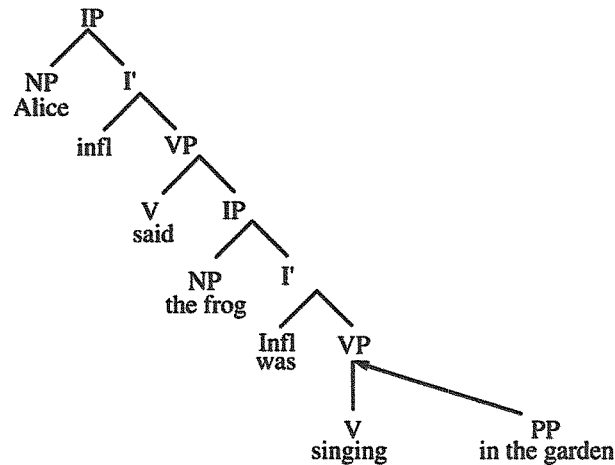
    c. # Joe looked the friend who had smashed his new car up.
    d. # I met the boy who Sam took to the park's friend.
    e. # The girl applied for the jobs that was attractive.

Our aim here is not to try to motivate the *existence* of the local attach-
ment preference. We will take that for granted. Nor, for the most part, do
we intend to choose among the many different versions of the local attach-

ment preference that have been proposed.[4] What we are more concerned with here is the question of how the local attachment preference interacts with other parsing biases, and in particular the question of how *strong* the local attachment preference is relative to other proposed structural factors in parsing.

As we will show, in asking about the strength of the local attachment preference we are faced with what seems at first to be a straightforward contradiction. First, there is some long-standing and well-known evidence that the local attachment preference is a weak bias in ambiguity resolution, and that there are a number of other biases that can overrule its choices. To contrast with that, we present the results of an experiment on a novel structural ambiguity which points to just the opposite conclusion, namely that the local attachment preference is rather stronger than has usually been assumed. In the last part of the paper, we show how it is possible to reconcile what seem to be opposing findings.

The evidence that local attachment is a relatively weak bias comes from examples like (2). As in (1) there are two possible attachments of the PP *in the garden*, but the choices are slightly different in this case. The alternatives are to make the PP an argument of the verb or a modifier right-adjoined to the object NP. Sentences (2c) to (2e) show further cases of V NP XP sequences in which XP could be attached to a projection of either the verb or the noun phrase.

Most versions of local attachment predict a preference for the NP-modifier reading of the PP *in the garden* in (2a), because this involves attachment to more recently built material. However, this prediction is incorrect, as both intuitions and a sizable body of experimental literature on the topic demonstrate (cf. Clifton, Speer, & Abney, 1991; Rayner, Carlson, & Frazier, 1983). Speakers reliably prefer the VP-attached interpretation of (2a).

---

[4] Among the family of proposals which we group under the heading local attachment preference we include *right association* (Kimball, 1973; Phillips, 1995), *late closure* (Frazier, 1978), *recency* (Gibson, 1991; Gibson et al., 1996), *attach low* (Abney, 1989), and *minimal connections* (Fodor & Frazier, 1983). These proposals differ with regard to whether locality is defined in terms of terminal strings, tree geometry, or grammatical attachment sites. The differences among these various versions of the local attachment preference are not important in this paper, except where noted.

(2) a. Alice put the singing frog in the garden.
   b.

```
                        IP
                       /  \
                     NP     I'
                   Alice   /  \
                         Infl   VP
                                / \
                               V   NP        PP
                              put  the singing frog   in the garden
```
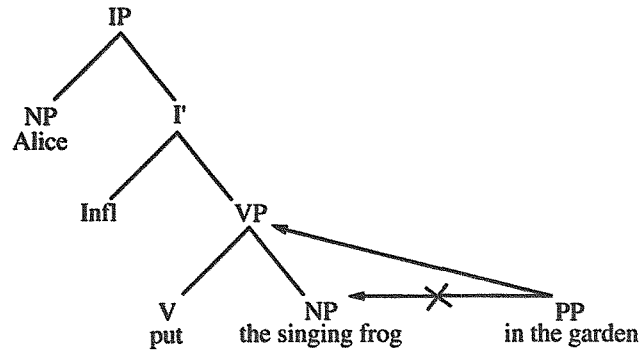
   c. # Joe bought the book for Susan to the party.
   d. # Alice saw the singing frog in the garden in the bathroom.
   e. # Henry told the intruder that he met to leave.

The conclusion that has typically been drawn from cases like those in
(2) is that there must be additional factors influencing the preferences speak-
ers show in examples like (2a) to (2e). Moreover, these other factors must
be *stronger* than local attachment. The literature contains a number of pro-
posals about what these other factors might be, including minimal attach-
ment (Frazier, 1978, 1987), argument attachment preferences (Abney, 1989;
Ford, Bresnan, & Kaplan, 1982; Gibson, 1991; Pritchett, 1988, 1992;
Schütze & Gibson, 1996), discourse accommodation (Altmann & Steedman,
1988; Crain & Steedman, 1985; Percus, 1995), and frequency (Spivey-
Knowlton & Sedivy, 1995).

The other piece of the locality puzzle, which is the main focus of this
paper, is that there is a further ambiguity which leads us to the opposite
conclusion from (2), namely that local attachment is *stronger* than the factors
listed in the preceding paragraph. This is the "matrix-relative" ambiguity
shown in (3) (cf. Gibson & Broihier, 1996). Given the incomplete sentence
because Rose praised the recipe I . . . , there are a couple of possible con-
tinuations. The NP *I* could be made either the subject of a *that*-less relative
clause modifying *the recipe*—this would mean that the clause that *I* is the
subject of should contain an object gap, as in *because Rose praised the
recipe I made for her birthday I also made it for her graduation*. Alterna-
tively, *I* could be attached as the subject of the matrix clause, in which case
the clause that *I* is the subject of should not contain an object gap, as in
because Rose praised the recipe I made it for her birthday. Further examples
of this ambiguity are given in the Appendix.

Although almost all parsing principles that we are aware of predict that subjects should always opt for the matrix clause attachment, the experiment described here shows evidence that subjects in fact opt for the relative clause attachment, which as far as we can tell *only* local attachment favors.

(3) a. Because Rose praised the recipe I made . . .

      i. . . . for her birthday I also made it for her graduation.

                                          (relative clause)

      ii. . . . it for her birthday.   (matrix clause)

b.
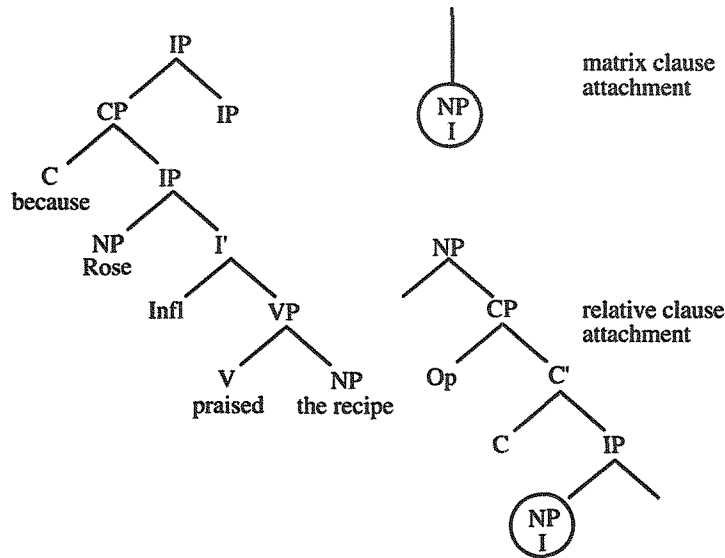
This preference leads us to believe that local attachment is a strong factor in ambiguity resolution. This conclusion clearly conflicts with what is usually concluded from the antilocality effects shown in (2). In the section Resolving the Locality Puzzle later on, we show how this conflict can be resolved.[5]

---

[5] See Cuetos and Mitchell (1988) and Gibson et al. (1996) for evidence that local attachment is outranked by another factor in parsing ambiguous Spanish NPs of the form *the daughter of the colonel who I met last week*. Spanish speakers show a (weak) preference for attaching the relative clause to the first NP in examples like this. Gibson et al. (1996) attributed this preference to the *predicate proximity* principle (see the section Resolving the Locality Puzzle below). We will have nothing to say about the English/Spanish contrast here.

COMPETING BIASES

Consider now what different well-known parsing principles predict about how the matrix-relative ambiguity will be resolved. The choice that the parser has to make is between a matrix clause attachment or an attachment in which the ambiguous NP is the subject of a relative clause modifying the object of the subordinate *because* clause, as shown in (3b) above. Notice that in the relative clause attachment rather more structure needs to be built in order to attach the overt NP.

First, it should be clear from looking at the alternative chunks of structure required for each of the two possible continuations in (3b) that Frazier's minimal attachment principle (Frazier, 1978, 1987) should prefer the matrix clause attachment. This is because the relative clause attachment requires more new structure to be built than the matrix clause attachment, under any way of counting new structural material.

Principles that favor attaching arguments or obligatory constituents over optional constituents (cf. Abney, 1989; Ford et al. 1982) should also choose the matrix clause attachment, since the matrix subject is an obligatory constituent in English, whereas a relative clause modifying the direct object of a subordinate clause is not at all obligatory.

Principles based on the idea that the parser prefers to leave as few predicate-argument relations unsatisfied as possible (e.g., Gibson, 1991; Pritchett 1988, 1992) should also favor the matrix clause attachment. If the ambiguous NP is attached as the matrix subject, then just the subject NP is lacking a theta-role assigner. If, on the other hand, the ambiguous NP is attached as the subject of an object relative clause, then two arguments will be lacking a theta-role assigner—both the subject of the relative clause and the null relativized object in SpecCP of the relative clause. Therefore, the matrix clause attachment is predicted to be preferred.

In addition, a selection relation holds between the *because* or *while* clause and the main clause. The simple fact that the subordinate clause requires the main clause in order to form a complete sentence could place additional pressure on the parser to pursue the matrix clause continuation.

Furthermore, in any theory in which the parser chooses the continuation which entails the simplest accommodation of the current discourse model (Altmann & Steedman, 1988; Crain & Steedman, 1985; Percus, 1995), the prediction is also that the matrix clause continuation should be chosen. This is because the relative clause implies the existence of some contrast set [e.g., of recipes in an example like (3)] which has not yet been inserted into the discourse model. The matrix clause attachment, on the other hand, entails no such unsupported implicatures.

These predictions are summarized in Table I. As is immediately clear, the only principle that favors the relative clause continuation for the matrix-relative ambiguity is the local attachment preference.

Thus, all parsing principles that make any choice about how the matrix-relative ambiguity in (3) should be resolved choose the matrix clause attachment, except for local attachment. The fact that *only* local attachment favors the relative clause continuation is useful, because it means that, if we find any evidence for a relative clause preference in resolving this ambiguity, then we may conclude that local attachment and none of the other factors listed is responsible for this preference. It would also show fairly unequivocally that local attachment is stronger than all of the other parsing principles listed in Table I, because that would be the only way of explaining how it could override the conflicting preferences that the principles in Table I predict to occur.

Note, however, that if we find that subjects show a matrix clause attachment preference, this is not very informative. Given that there are so many different reasons to prefer the matrix clause attachment over the relative clause attachment, it would be hard to know which factor(s) is (are) responsible for the parser's choice of the matrix clause attachment.

EXPERIMENTAL EVIDENCE

We conducted an experiment to test the intuitive preference for relative clause attachment in the matrix-relative ambiguity. The predictions of local attachment and other proposed parsing strategies were discussed in the previous section.

Methods

Subjects. Forty-seven native English speakers from MIT (primarily undergraduate students) participated, for $8.00 each.

Materials. There were four conditions: matrix clause and relative clause resolutions, with ambiguous and unambiguous versions of each. The difference between the matrix and relative clause conditions was that the relative clause conditions contained an object gap where the matrix clause conditions contained an overt pronoun. In the unambiguous relative clause condition, the relative clause was introduced by the complementizer *that*. In the unambiguous matrix clause condition the direct object NP in the subordinate clause contained a possessor, e.g., *the recipe* was replaced by *my recipe*. The relevance of this is that possessors block relativization, so that when the following (ambiguous) NP is encountered the parser knows that it must

Table I. Factors Potentially Contributing to Matrix and Relative Clause Attachments

Factor	Preference	Reason
Minimal attachment	Matrix	Fewer nodes required to build matrix Spec, IP than CP complex plus Spec, IP in relative clause.
Attach arguments/obligatory constituents	Matrix	Matrix subject is obligatory, relative clause is not.
Complete predicate-argument relations	Matrix	Matrix attachment has one argument missing a predicate; relative clause attachment has two arguments missing a predicate.
Semantic completeness	Matrix	*Because* or *while* clause needs an antecedent matrix clause; relative clause does nothing for this.
Discourse accommodation	Matrix	In null context there's no contrast set for *the recipe*, so no pressure to modify.
Tense-matching constraint (see below)	Matrix	Increased matrix clause activation due to tense-parallel imposed by *while, after, until* clauses.
Local attachment, right associ-ation, recency	Relative	Closer association to most recently attached lexical material.

be the matrix subject, because the relative clause continuation is not a possible option.[6]

(5) a. *Relative clause, ambiguous (RA)*
      Because Rose praised the recipe I made for her birthday it was worth all the effort.
   b. *Relative clause, unambiguous (RU)*
      Because Rose praised the recipe that I made for her birthday it was worth all the effort.
   c. *Matrix clause, ambiguous (MA)*
      Because Rose praised the recipe I made it for her birthday as a surprise.
   d. *Matrix clause, unambiguous (MU)*
      Because Rose praised my recipe I made it for her birthday as a surprise.

[6] It is not strictly true that NPs containing a possessor do not allow relativization (thanks to Gregory Ward for pointing this out). However, the combination of the fact that relative clauses following NPs containing possessors are generally interpreted as non-restrictive and the fact that *that*-less relative clauses must receive a restrictive reading means that relative clause readings of examples like our (5d) should be either impossible or extremely marked. Furthermore, the results of the experiment imply that the possessors in the (5d) condition were indeed an effective disambiguator.

Two additional properties varied in our items. Half of the items described a *nontemporal* relation, using complementizers like *because* [cf. Example (5)], *since*, or *although*, and the other half described a *temporal* relation, using complementizers like *while, when*, or *after*. In addition, the ambiguous NP was a pronoun in 11 of the items and a full NP in the 13 remaining items [e.g., *While I talked with the lawyer John was watching (him) at the party . . .*]. A complete list of materials is given in the Appendix.

We should point out that although the factor ±temporal was balanced in our materials, it was not varied independently of the factor ±pronominal. Ten of the eleven items in which the ambiguous subject NP was a pronoun occurred in items describing a nontemporal relation. This fact becomes relevant below.

Twenty-four items each with four forms like those shown in (5) were constructed. The 24 experimental items were combined with 90 fillers to form four lists. The fillers were of approximately the same length and complexity as the experimental items. The experimental items were counterbalanced across the lists so that each list contained six items from each condition and exactly one version of every item. Ten practice items were also constructed to be similar to the fillers.

Procedure. The stimuli were presented on a Macintosh Centris computer using a word-by-word self-paced moving-window paradigm (Just, Carpenter, & Wooley, 1982). A subject saw two screens of instructions, followed by 10 practice items and then the 114 experimental and filler items. The experimental and filler items were presented in a different random order for each subject. Each sentence was followed by a question, to ensure that subjects had understood the sentences. All trials on which the question was incorrectly answered were excluded prior to further analysis. One subject was excluded because of an extremely high error rate in the questions.

Results

A regression equation predicting reading time from word length was constructed for each subject, using all items (filler and experimental). At each word, the reading time predicted by the subject's regression equation was subtracted from the actual measured reading time, and all analyses were performed on these differences (residual reading times). This transformation removes extraneous variance by subtracting out a baseline for each subject, and by controlling for noise due to length effects (Ferreira & Clifton, 1986; Trueswell & Tanenhaus, 1991).

Trials on which the subject answered the comprehension question incorrectly were excluded from all of the following analyses. This affected 9% of trials. One subject for whom data were not available for all conditions

was also excluded. In addition, all trials with residual reading times greater than 1,000 ms were also excluded from further analysis. This affected less than 1% of the trials.

No differences were found preceding the point of disambiguation (object pronoun in matrix clause conditions, PP in relative clause conditions).

On the four-word region beginning at the point of disambiguation, there were main effects of ambiguity, attachment site, and connective (i.e, ±temporal), as follows: ambiguous = 1.30 ms, unambiguous = -18.44 ms, $F_1(1, 45) = 8.15$, $p < .01$, $F_2(1, 11) = 10.2$, $p < .005$; matrix = -19.55 ms, relative = 2.42 ms, $F_1(1, 45) = 12.3$, $p < .01$, $F_2(1, 11) = 7.06$, $p < .05$; temporal = -1.58 ms, nontemporal = -15.55 ms, $F_1(1, 45) = 6.06$, $p < .05$, $F_2(1, 11) = 2.08$, $p = .16$. The main effect of connective did not reach significance by items. There was also an Ambiguity × Attachment Site × Connective Interaction, shown in Fig. 1, which we focus on in what follows. This interaction was only marginally significant by items: $F_1(1, 45) = 4.99$, $p < .05$, $F_2(1, 11) = 3.01$, $p < .1$. The nontemporal examples show evidence for a relative clause attachment preference, whereas the temporal examples show evidence for a matrix clause attachment preference.[7]

Figure 2 compares residual reading times for ambiguous and unambiguous versions of the matrix clause conditions with the nontemporal items. There is a highly significant slowdown in the ambiguous condition, which begins as soon as the disambiguating pronoun is read, and extends into the first two words of the following prepositional phrase. Region 3: ambiguous = 55.45 ms, unambiguous = -9.57 ms; $F_1(1, 45) = 8.71$, $p < .01$; $F_2(1, 11) = 8.30$, $p < .05$. Region 4: ambiguous = 4.78 ms, unambiguous = -37.7 ms; $F_1(1, 45) = 5.85$, $p < 0.05$; $F_2(1, 11) = 7.98$, $p < .05$. Region 5: ambiguous = -11.78 ms, unambiguous = -48.38 ms; $F_1(1, 45) = 5.24$, $p < .05$; $F_2(1, 11) = 7.81$, $p < .05$. Regions 3 to 5 combined: ambiguous = 16.24 ms, unambiguous = -31.48 ms; $F_1(1, 45) = 12.1$, $p < .01$; $F_2(1, 11) = 11.55$, $p < .01$. This is a classic "filled-gap" effect (cf. Stowe, 1986). No other regions showed significant effects of ambiguity. In the nontemporal relative clause conditions, the ambiguous items showed a small slowdown at the PP following the gap, but it did not reach significance or marginal significance on any individual word or the PP as a whole, in clear contrast to the matrix clause conditions.

[7] This four-word region was different in the matrix and relative clause conditions. For example, in an example like (3a) the four-word region for the matrix clause conditions is *it for her birthday*, and the four-word region for the relative clause condition is *for her birthday as*. We cannot exclude the possibility that this difference contributed to the observed main effect of attachment site; however, given that it is the Ambiguity × Attachment Site × Connective interaction that we focus on in what follows, this issue does not affect our conclusions.
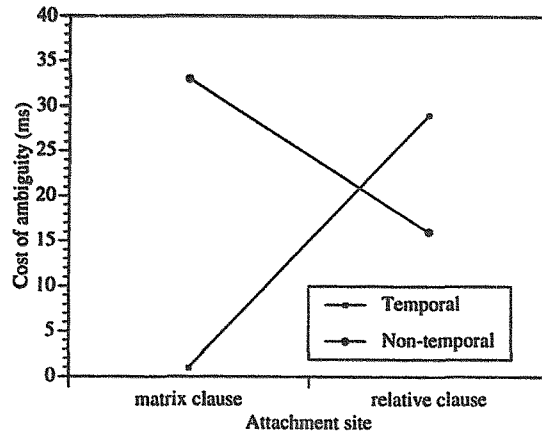
Fig. 1. Ambiguity × Site × Connective interaction at four-word region following disambiguation.

The items describing a temporal relation showed the opposite pattern of results. The matrix clause conditions showed no effect of slowdown due to ambiguity at any word or region, whereas in the relative clause conditions we found a strong garden path effect, starting at the second word of the disambiguating PP and extending for a region of three to four words, as Fig.
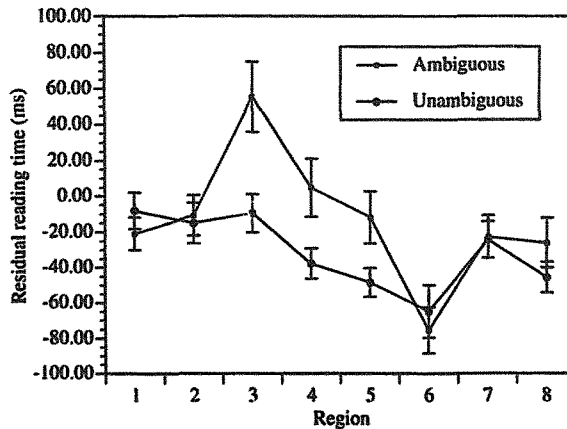
Fig. 2. Residual reading times in matrix/−temporal conditions.

3 shows. Region 5: ambiguous = 16.72 ms, unambiguous = −21.73 ms; $F_1(1, 45) = 4.57, p < .05$; $F_2(1, 11) = 6.41, p < .05$. Region 6: ambiguous = 23.84 ms, unambiguous = −4.82 ms; $F_1(1, 45) = 1.4, p < .3$; $F_2(1, 11) = 0.49, p = .5$. Region 7: ambiguous = 55.71 ms, unambiguous = −2.64 ms; $F_1(1, 45) = 7.82, p < .01$; $F_2(1, 11) = 9.84, p < .01$. This garden-path effect is not as immediate as in the nontemporal matrix clause conditions, where it occurred immediately at the disambiguating pronoun. This slight delay is not surprising, though, given that disambiguation in these conditions occurs when the subject notices the *absence* of an overt object NP in the relative clause conditions, whereas it only requires noticing the *presence* of an overt object NP in the matrix clause conditions. Also, given the possibility of heavy NP shift to the right of the PP, the absence of an NP between the verb and the PP does not provide definitive evidence for a relative clause analysis.

Combining data for all sites and connectives, at the first word following the PP (i.e., Region 7) there was a marginally significant effect of ambiguity: ambiguous = 3.33 ms, unambiguous = −11.0 ms; $F_1(1, 45) = 3.44, p < .1$; $F_2(1, 22) = 3.88, p < .1$. There was a highly significant effect of attachment site: matrix = −30.51, relative = 22.85; $F_1(1, 45) = 25.66, p < .001$; $F_2(1, 22) = 23.65 p < .001$. The following interactions were also significant by subjects, items, or both: Ambiguity × Attachment Site, $F_1(1, 45) = 6.28, p < .05$; $F_2(1, 22) = 5.26, p < .05$; Ambiguity × Connective, $F_1(1, 45) = 1.36, p = .25$; $F_2(1, 22) = 4.66, p < .05$; Ambiguity × Attachment Site × Connective, $F_1(1, 45) = 3.04, p < .1$; $F_2(1, 22) = 5.76, p < .05$.
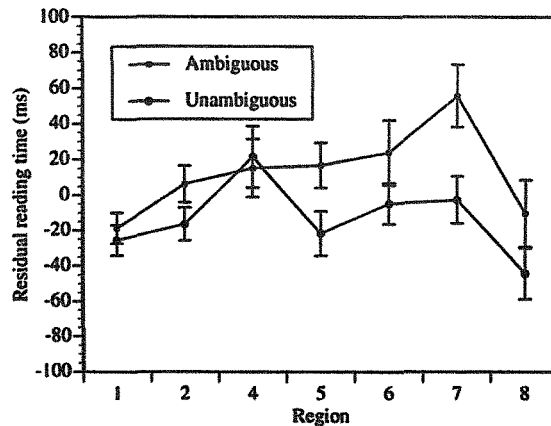

Fig. 3. Residual reading times in relative/temporal conditions.

At the following word (i.e., Region 8), there was a main effect of ambiguity, which was significant by subjects and marginally significant by items: ambiguous $= -24.97$, unambiguous $= -45.47$; $F_1(1, 45) = 6.56$, $p < .05$; $F_2(1, 22) = 4.04$, $p < .1$. No other comparisons were significant.

Discussion

Most importantly, the fact that in the nontemporal conditions we found a preference for the relative clause attachment confirms that local attachment is an extremely strong structural factor in parsing. It also shows that local attachment is strong enough to override the opposing forces to follow the matrix clause attachment that the other principles listed in Table I predict.

A comment is in order at this point on the issue of punctuation. Our stimuli were presented with no punctuation, and it has been suggested to us that this may be a confounding factor in our results. The objection usually goes as follows: In written text the matrix clause attachments would often be disambiguated by means of a comma. Therefore the absence of a comma in these stimuli could be responsible for leading people to initially pursue the relative clause attachment. There are a couple of reasons why we think this is an unlikely explanation for the results.

First, as the split in our results between temporal and nontemporal items shows, speakers were able to pursue the matrix clause attachment in the absence of a comma. In written text the subordinate clause and the matrix clause are often not separated by a comma. Second, we agree that the role of a comma when it occurs in examples like these is to aid comprehension, but it is important to ask why the comma should ever be necessary in these cases. The answer is presumably that the comma is used to counteract the local attachment preference for the relative clause parse. Therefore, the fact that commas are sometimes used in examples like ours confirms rather than confounds our claim that there is a local attachment preference in these sentences (cf. Frazier & Rayner, 1988 for a similar argument).

We owe an explanation for why the contrast between the temporal and the nontemporal items should have led to such a clear difference in parsing preferences. As pointed out above, the factors ±temporal and ±pronominal were not independent in our materials, so further experiments will be required in order to determine whether one or both of these factors was responsible for the split in the results. We can, however, offer the following possible explanations.

Tense Matching. If the factor ±temporal was responsible for the split in the results, then this may be due to the fact that the temporal items observed a tense matching requirement between the subordinate and the matrix clauses, but the nontemporal items did not.

As (6a) and (6b) show, when the two clauses describe a temporal relation, the choice of tense in the first clause tightly constrains the choice of tense in the second clause. In (6a) for example, the subordinate clause *while John was eating his lunch*, which contains a past tense verb, can be followed by a past tense matrix clause like *he was watching TV*. But if we replace this with a present tense or future tense main clause, the sentence becomes ungrammatical. Sentences describing temporal relations are therefore subject to a *tense-matching constraint*.

(6) a. While John was eating his lunch . . .
        . . . he was watching TV   past
        * . . . he watches TV   * present
        * . . . he will be watching TV   * future
  b. After Mary gets off the bus on Mondays . . .
        * . . . she bought a cup of coffee   * past
        . . . she buys a cup of coffee   present
        * . . . she will buy a cup of coffee   * future

In (7a) and (7b), though, which give examples of nontemporal sentences, no tense-matching constraint is operative. As long as general constraints on tense sequencing in English are respected, more or less any tense can combine with more or any other tense. In (7b) for example, the subordinate clause *although Helen drives a microbus* can be followed by past, present, or future main clauses with no difficulty.

(7) a. Because John was eating a greasy sandwich at his desk . . .
        . . . he was in trouble with the boss   past
        . . . he is in trouble with the boss   present
        . . . he will be in trouble with the boss   future
  b. Although Helen drives a VW microbus . . .
        . . . she used to own a Cadillac   past
        . . . she also owns a Cadillac   present
        . . . she will soon own a Cadillac   future

To see why the tense-matching constraint should make a difference to the resolution of the matrix-relative ambiguity, imagine that the parser is working its way through a subordinate clause headed by a temporal complementizer like *while*. We assume that the tense-matching constraint is immediately active during parsing. So as soon as the complementizer *while* is encountered, the parser knows to expect both a subordinate clause and a matrix clause and it also knows that the two clauses will match in tense. Therefore, as soon as the tensed verb is reached, the parser is immediately in a position to build tense features into the Infl node of the matrix clause, and so these features are built right away.

The relevance of this is that by working on building the matrix inflection, the matrix IP projection is activated to a certain degree. The effect of this when the parser encounters the ambiguous NP is to add to the other factors that are lobbying for a matrix clause attachment, and this turns out to be just enough to override the strong pull for relative clause attachment that comes from the local attachment preference. In the nontemporal cases, on the other hand, the absence of the tense-matching constraint means that local attachment is able to override any biases toward matrix clause attachment.[8]

Pronominals Versus Full NPs. Another possible reason for the split in our results involves the contrast between sentences in which the ambiguous NP is a pronoun and sentences in which it is a full NP (e.g., *the janitor*). Recall that 10 of the 12 sentences describing a nontemporal relation (relative clause preference) had a pronoun as the ambiguous NP, and 11 of the sentences describing a temporal relation had a full NP as the ambiguous NP (matrix clause preference). In almost all cases the pronouns had an antecedent in the subordinate clause, and the full NPs always introduced novel discourse referents.

The full NPs may be preferentially attached in the matrix clause because of discourse factors. All of the ambiguous relative clauses lack a relative pronoun, and are therefore restrictive relative clauses. In a felicitous discourse, the content of a restrictive relative clause is given information rather than new information. But the fact that the full NPs pick out novel discourse referents entails that the relative clauses cannot contain given information alone. This may provide sufficient pressure to avoid the local attachment and pursue the matrix clause attachment. In the examples with ambiguous pronouns, on the other hand, the first word of the relative clause

[8] Townsend and Bever (1978) studied the effect of a range of temporal and causal connectives on subjects' speed of comprehending embedded clauses, using a task in which subjects were asked to rate the consistency of a verb-object probe with the meaning of a previously presented embedded sentence. They argued that embedded clauses expressing causes are more fully interpreted more quickly, and clauses expressing an "adversative" relation (e.g., *though*) are interpreted less fully or less quickly. Temporal connectives fall between these two extremes. Our finding about the effect of ± temporal is probably independent of Townsend and Bever's results, because the range of connectives in our +temporal and −temporal sets of stimuli span equal ranges on Townsend and Bever's causal-adversative continuum.

picks out an existing discourse referent, so the same "discourse penalty" is not incurred if the relative clause attachment is pursued.[9]

This gives just a sketch of why the factors ±temporal or ±pronominal could lead to the bifurcation that we observe in our results. Further experiments are underway which attempt to separate these two factors, by independently manipulating ±pronominal and ±temporal.

We have assumed here that the relative clause attachment preference in the −temporal conditions is due to a strong local attachment preference, and that the opposite attachment preference in the +temporal conditions is due to the additional effect of a matrix-attachment enhancing factor. An alternative possibility is that the matrix attachment preference is the more basic finding, and that there is a low-attachment enhancing factor which causes the results in the −temporal conditions. However, we are unaware of any factor which might lead to a stronger bias for local attachment in nontemporal conditions than temporal conditions.

There is one discourse-based factor which potentially favors relative clause attachments, but this factor should favor relative clause attachments across the board rather than in −temporal conditions only. Notice that the direct object in the initial clause is definite in all stimuli. Introduction of this definite may cause a presupposition violation, because the definite is not already established in the discourse. Treating the following material as a relative clause modifying the definite could be a way of lessening or avoiding this presupposition violation. This amounts to the opposite of the claim of Crain and Steedman (1985) and Altmann and Steedman (1988) that relative clauses modifying definite NPs are avoided because they violate discourse conditions (see Percus, 1995, for discussion of conditions under which modifiers are and are not felicitous).

However, if this particular discourse factor were responsible for the low attachment with the −temporal items, we would lack an explanation for why low attachment was also not found with the +temporal items.

RESOLVING THE LOCALITY PUZZLE

We are now in a position to return to the locality puzzle that we began with. Recall what the problem is: The results of our experiment lead to the

---

[9] This account of the effect of the ±pronominal factor is related to but different from proposals of Crain and Steedman (1985) and Altmann and Steedman (1988). For these authors, building a relative clause carries a cost when it presupposes a contrast that is not supported in the prior discourse. What we are assuming here is that building a relative clause does not inherently carry a cost, but it does carry a cost if its referents give rise to unsupported presuppositions.

conclusion that the local attachment preference is strong, and can override a variety of other potential factors in parsing. But this has to be reconciled with the antilocality effects that we pointed out at the beginning, and that are motivated by a sizable body of work. The key evidence is summarized in (8).

(8) a. # Because Rose liked the recipe I made it for her birthday.

          (Local attachment > other factors in Table I)

    b. # Alice put the singing frog in the garden in the circus.

          (Local attachment < other factors in Table I)

We can see a couple of ways of resolving the locality puzzle, each of which we have pursued in other work. The first approach focuses on the fact that the locality puzzle only arises if it is assumed that parsing biases are ranked, so that one factor will always win out over another. If the strict hierarchy view is dropped and a series of weaker factors is allowed to "gang up on" a stronger factor, then it is fairly easy to solve the locality puzzle. The second approach questions the assumption that the VP attachment preferences that are taken to show antilocality effects are really antilocality effects. A reexamination of the syntax of these examples shows that they are actually consistent with a local attachment preference. If this is the case, then the locality puzzle does not arise either.

Examples (9) to (11) show how the puzzle is resolved under an approach in which different structural constraints are weighted and can conspire (cf. Gibson, 1991; Gibson & Broihier, 1996; Gibson et al., 1996). The relevant constraints are shown in (9). The strongest constraint, recency, favors attachments to recently built structure; the next strongest constraint places a cost on parses in which there are arguments whose theta-role assigner has not yet been determined; finally predicate proximity favors attachments to projections of verbs and other predicates.

(9) a. Recency favors attachments to projections of recent items.
    b. Local theta violations are incurred by arguments lacking a theta-role.
    c. Predicate proximity favors attachment as structurally close to a predicate as possible.

In the matrix-relative ambiguity shown in (10), the relative clause continuation is chosen, because one recency violation and one local theta violation are worse than two local theta violations. Neither representation is associated with a predicate proximity violation, because the attachment of the subject NP *I* is equally close to a predicate in each. In both cases, this NP is the subject of a predicate VP to come.

(10) # Because Rose liked the recipe I made it for her birthday.

   ☞ Relative clause attachment   0 recency violations
                                                 2 local theta violations
                                               (subject and relativized object)
                                               0 predicate proximity violations
      Matrix clause attachment   1 recency violation
                                               1 local theta violation
                                             0 predicate proximity violations

It follows from other results that a recency violation is associated with greater cost than is a theta violation (Gibson, 1991).[10] Thus the preferred attachment according to this theory is the more local relative clause attachment, in spite of its two theta violations.

Turning now to the antilocality effect repeated in (11), notice that if we consider only recency and local theta violations this theory incorrectly predicts that the PP will be attached to the NP rather than to the verb. This is because the matrix-relative ambiguity showed that recency violations must incur a greater cost than local theta violations. However, this is one of the situations where predicate proximity plays a role: By conspiring with the theta-attachment constraint it is able to override the strength of the recency preference.

(11) # Alice put the singing frog in the garden in the circus.

      Modifier attachment   0 recency violations
                                               1 local theta violation
                                             1 predicate proximity violation
   ☞ Argument attachment   1 recency violation
                                             0 local theta violations
                                             0 predicate proximity violations

---

[10] Specifically, a difference of a single recency violation is enough to cause a strong preference for the more local attachment, as in (i):

    (i) The teaching assistant told the professor that the students were confused during the class.

The PP *during the class* can ambiguously attach to either the matrix IP or the embedded IP. There is a single recency violation difference between the two attachments, but there are no other cost differences. The preference for low attachment is strong enough that semantic factors cannot override the attachment preference initially, as evidenced by the difficulty associated with the processing of (ii):

    (ii) # The teaching assistant told the professor that the students will be confused yesterday during the class.
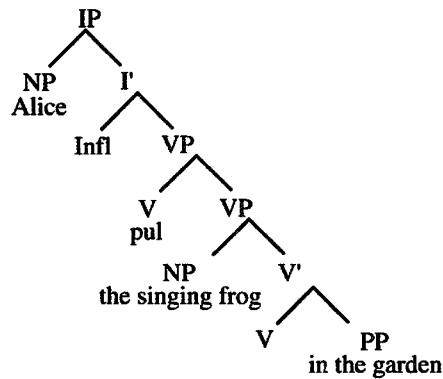
In contrast, a difference of a single theta violation does not cause as such a strong preference that semantic and pragmatic factors cannot override the thematic preference. See Gibson (1991) and Gibson, Hickok, and Schütze (1994) for more details.

Under this view, then, the locality puzzle involves a genuine conflict, but such conflicts are the hallmark of this approach.

The second approach to the locality puzzle, argued for in Phillips (1995) (building in part on Fodor & Frazier 1983), claims that there is really no conflict between the locality and antilocality effects. What has been thought of as an antilocality effect is in fact entirely consistent with local attachment.

A body of work on constituent structure from the last 10 years argues that VP-modifier phrases and the second argument in double complement structures are actually sisters of the verb in a right-branching VP structure like (12) (cf. Larson, 1988, 1990; Pesetsky, 1995; Phillips, 1996; Stroik, 1990).[11]

(12)

```
                IP
              /    \
            NP      I'
           Alice   /  \
               Infl    VP
                      /   \
                     V     VP
                    pul   /   \
                        NP     V'
             the singing frog  / \
                              V   PP
                           in the garden
```

These syntactic results become relevant if the local attachment is instantiated by the principle of *right association* (Kimball, 1973, Phillips, 1995, 1996), a principle which favors construction of right-branching structures.[12] As (13a) and (13b) show, both the matrix-relative ambiguity and the

---

[11] The evidence for this involves the results of tests of coordination, binding, idiom-formation, polarity item licensing, etc., in conjunction with the assumption that these tests transparently diagnose constituency or c-command relations. See Phillips (1996) for justification of this assumption, which has often been doubted in the literature.

[12] Maximally right-branching structures are structures in which precedence relations and c-command relations among pairs of terminal elements are identical, i.e., X precedes Y iff X c-commands Y. Therefore, we can talk of one structuring of a set of elements being *more right branching* than another structuring of the same set of elements if there are fewer mismatches between precedence and c-command relations.

V-NP-PP ambiguities are resolved in favor of the more right-branching alternative, and therefore conform to the predictions of right association.[13]

(13) a. # Because Rose liked the recipe I made it for her birthday.

    ☞ Relative clause attachment   more right-branching continuation

    Matrix clause attachment   less right-branching continuation

  b. # Alice put the singing frog in the garden in the circus.

    Modifier attachment   less right-branching continuation (adjunction to NP)

    ☞ Argument attachment   more right-branching continuation (sister of V)

We do not intend to choose between these two alternatives here. Nor is the choice among them important from the perspective of this paper, because our main points, which are consistent with both alternatives, are that the local attachment preference is a strong structural factor in parsing and that biases must be weighted rather than ranked.

CONCLUSIONS

Based on our experiment on the matrix-relative ambiguity, we hope to have shown that the local attachment preference is stronger than has often been assumed. In particular, its choices are able to override each of the other potential factors listed in Table I.

Our results showed a split in attachment preferences between sentences describing temporal relations and sentences describing nontemporal relations. We suggested an explanation for this split, based on the effects of a tense-matching constraint, and an alternative explanation based on whether the ambiguous word is a pronoun or a full NP.

Although our findings regarding the strength of the local attachment preference may appear to conflict with a body of well-known evidence that local attachment is not so strong, we have shown a couple of ways in which the "antilocality" results can be reconciled with our results.

---

[13] Note that the kind of antilocality effects involving complex NPs of the kind *the daughter of the colonel who . . .* reported for Spanish by Cuetos and Mitchell (1988) and for Dutch by Brysbaert and Mitchell (1995) is not accounted for under this proposal.

APPENDIX: MATERIALS FOR EXPERIMENT

In each item the pronoun in parentheses was present in the matrix clause and not the relative clause conditions. The two phrases enclosed in parentheses at the end of each sentence are the completions of the relative clause and matrix clause conditions, respectively.

Temporal, −Pronominal

1. While I talked with the lawyer John was watching (him) at the party (I became rather nervous/and that made me rather nervous).
2. Until I trained the dog everybody was avoiding (it) like the plague (I had very few visitors/because they were afraid of getting bitten).
3. After I watered the plant the housekeeper neglected (it) for many months (it slowly recovered/and it died).
4. As the president outlined the speech his advisor drafted (it) for the convention (he realized how difficult his task would be/that was fast approaching).
5. As the king's army lost control of the castle the enemy was destroying (it) with huge cannons (some reinforcements arrived/and closing in on where the king hid).
6. When the collector displayed the painting the expert identified (it) as a fraud (the police were contacted/within minutes).
7. When the chef served the dessert the guests liked (it) quite a lot (the treaty was quickly settled/and told him so).
8. When the boys discovered the canoe their dad was fixing (it) in the garage (they knew what they were getting for Christmas/as a Christmas present).
9. By the time the girls found the gifts their mother had wrapped (them) for the party (they were quite late/and it was time to leave).
10. When the company board promoted the accountant John disliked (him) for his success (John resigned/and his wealth).
11. As the residents looked at the building the crane was demolishing (it) with a wrecking-ball (some of them couldn't help applauding/and leaving only rubble).

Temporal, +Pronominal

12. When Mike arrived at the house I described (it) in great detail (he was quite impressed/because he was interested in old building styles).

Nontemporal, −Pronominal

13. Because the senator proposed the bill the speaker opposed (it) in the house (he resolved to unseat the speaker in the election/and attacked the senator on TV).
14. Since the author's friends bought all the books the publisher distributed (them) to major booksellers (the novel made the best-seller lists/and made enormous profits).

Nontemporal, +Pronominal

15. Because Rose praised the recipe I made (it) for her birthday (it was worth all the effort/as a surprise).
16. Although the boy was afraid of the dog he annoyed (it) in the park (he tried to remain calm/with his friends).
17. Because Joe liked the children he saw (them) on the weekends (he found it hard to leave/whenever he could).
18. Although I liked the flowers the janitor removed (them) from the office (I didn't get angry/while I was away).
19. Because Fred was unhappy about the car he sold (it) to the dealer at half price (he never bought the same make of car again/to get rid of it).
20. Although the dissident eventually destroyed the letters he had kept (them) beneath the floorboards (it pained him to do so/for months).
21. Although Gwen and Phil were proud of the vase they had obtained (it) at a garage sale (it wasn't really valuable/sale for five dollars).
22. Because the police didn't trust the witness they had interrogated (him) for five hours (they ran a lie detector test/before the trial).
23. Since the students had enjoyed the classes they completed (them) with no trouble (they encouraged their friends to take the same classes/and got good grades).
24. Because the conference organizers liked the caterers they hired (them) for the banquet (they asked them back the following year/in honor of the former president).

REFERENCES

Abney, S. (1989). A computational model of human parsing. *Journal of Psycholinguistic Research, 18,* 129–144.

Altmann, G., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition, 30,* 191–238.

Brysbaert, M., & Mitchell, D. (1996). Modifier attachment in sentence parsing: Evidence from Dutch. *Quarterly Journal of Experimental Psychology, 49,* 664–695.

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

Crain, S., & Steedman, M. (1985). On not being led up the garden path: the use of context by the psychological parser. In D. Dowty, L. Karttunen, & A. Zwicky (Eds.), *Natural language parsing: Psychological, computational and theoretical perspectives.* Cambridge, UK: Cambridge University Press.

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

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

Fodor, J. D., & Frazier, L. (1983). *Local attachment in a one-stage parser.* Unpublished manuscript.

Ford, M., Bresnan, J. & Kaplan, R. (1982). A competence based theory of syntactic closure. In J. Bresnan (Ed.), *The mental representation of grammatical relations* (pp. 727–796). Cambridge, MA: MIT Press.

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

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII.* Hillsdale, NJ: Erlbaum.

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

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

Gibson, E., & Broihier, K. (In press). Optimality theory and human sentence processing. In P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis, D. Peselsky (Eds), *Is the best good enough?* Cambridge, MA: MIT Press.

Gibson, E., Hickok, G., & Schütze, C. (1994). Processing empty categories in a parallel parsing framework. *Journal of Psycholinguistic Research, 23,* 381–405.

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

Just, M. A., Carpenter, P. A., & Wooley, 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.

Larson, R. (1988). On the double object construction. *Linguistic Inquiry, 19,* 335–391.

Larson, R. (1990). Double objects revisited: Reply to Jackendoff. *Linguistic Inquiry, 21,* 589–632.

Percus, O. (1995). A horse raced to logical form. In C. Schütze, J. Ganger, K. Broihier (Eds), *Papers on Language Processing and Acquisition* (pp. 473–519). (MIT Working Papers in Linguistics 26). Cambridge, MA: MITWPL.

Pesetsky, D. (1995). *Zero syntax.* Cambridge, MA: MIT Press.

Phillips, C. (1995). Right association in parsing and grammar. In C. Schütze, J. Ganger, K. Broihier (Eds), *Papers on Language Processing and Acquisition* (pp. 37–93) (MIT Working Papers in Linguistics 26). Cambridge, MA: MITWPL.

Phillips, C. (1996). *Order and structure*. Unpublished doctoral dissertation, MIT.

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

Pritchett, B. (1992). *Grammatical competence and parsing performance*. Chicago, IL: University of Chicago Press.

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

Schütze, C., & E. Gibson. (1996, March). *Argumenthood and PP-attachment*. Poster presented at the Ninth Annual CUNY Conference on Sentence Processing, New York.

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

Stowe, L. (1986). Evidence for on-line gap location. *Language and Cognitive Processes, 1*, 227–245.

Stroik, T. (1990). Adverbs as V-sisters. *Linguistic Inquiry, 21*, 654–661.

Townsend, D., & Bever, T. (1978). Interclause relations and clausal processing. *Journal of Verbal Learning and Verbal Behavior, 17*, 509–521.

Trueswell, J. C., & Tanenhaus, M. K. (1991). Tense, temporal context and syntactic ambiguity resolution. *Language and Cognitive Processes, 6*, 303–338.