LANGUAGE AND COGNITIVE PROCESSES, 2003, *18* (4), 469–503

Evidence against the use of subcategorisation frequency in the processing of unbounded dependencies

Martin J. Pickering
Department of Psychology, University of Edinburgh, Edinburgh, UK

Matthew J. Traxler
Department of Psychology, University of South Carolina, SC, USA

Three experiments investigated strategies readers use to process locally ambiguous unbounded dependency constructions. Using self-paced reading, Experiments 1 and 2 manipulated subcategorisation preferences and plausibility, and found that readers initially misanalysed a noun-phrase filler as an object of a verb, whether or not the verb preferentially took a noun-phrase object. Using eye-movement monitoring, Experiment 3 showed that readers misanalysed when the verb preferentially did not take a noun-phrase object in normal reading. These results indicate that readers do not select an initial analysis for an unbounded dependency construction on the basis of subcategorisation frequency. Indeed, the preference for forming such dependencies as quickly as possible does not seem to be affected by the verb's subcategorisation preferences. We discuss the implications of the results for theories of parsing.

The role of subscategorisation preferences is a central concern of current theories of parsing. In recent years, two broad classes of theories have competed to explain available data. One class assigns a central role to frequency and assumes that the processor can employ all potentially

Requests for reprints should be addressed to Martin Pickering, Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh EH8 9LW, UK.
Email: martin. pickering@ed.ac.uk

The order of authorship is arbitrary. We acknowledge Roger van Gompel and the Sentence Processing Group at the Universities of Glasgow and Edinburgh for their advice and encouragement. The research was also supported in part by NIH Grant No. 1 R01 HD 040865-01A1 (both awarded to Matthew J. Traxler) and by a University of Edinburgh Development Trust grant (awarded to Martin J. Pickering).

http://www.tandf.co.uk/journals/pp/01690965.html DOI: 10.1080/01690960344000017

relevant information in making initial decisions about syntactic analysis. The most influential accounts of this type are constraint-based theories, which assume that the processor computes different analyses in parallel, and chooses between them using all relevant information in a single stage (Garnsey, Pearlmutter, Myers, & Lotocky, 1997; MacDonald, 1994; MacDonald, Pearlmutter, & Seidenberg, 1994; McRae, Spivey-Knowlton, & Tanenhaus, 1998; Spivey-Knowlton & Sedivy, 1995; Trueswell, Tanenhaus, & Garnsey, 1994; Trueswell, Tanenhaus, & Kello, 1993, cf. Taraban & McClelland, 1988; Tyler & Marslen-Wilson, 1977). The processor immediately uses subcategorisation-preference information to favour the analysis that employs the most frequent subcategorisation (all other constraints being equal). Similar predictions follow from accounts in which initial parsing preferences are directly determined by frequency of exposure to particular subcategorisation frames for a given verb (Ford, Bresnan, & Kaplan, 1982) or frequency of exposure to similar syntactic structures (Mitchell, Cuetos, Corley, & Brysbaert, 1995).

However, other accounts do not assume that the processor favours the analysis with the most frequent subcategorisation. Most of these accounts assume that the processor ignores frequency, along with other sources of potentially relevant information, in its initial decisions about syntactic analysis. The best-known such account is the Garden Path theory, in which the processor's initial choice of analysis is based on general syntax-driven strategies like Minimal Attachment, which says, essentially, choose the analysis requiring the postulation of the fewest nodes in a phrase structure tree (Ferreira & Clifton, 1986; Ferreira & Henderson, 1990; Frazier, 1979, 1987; Rayner, Carlson, & Frazier, 1983, cf. Abney, 1989; Crocker, 1996; Gorrell, 1995; Pritchett, 1992). In such accounts, subcategorisation preferences are ignored during initial analysis (as are other factors such as plausibility), though they may well play a role in reanalysis.

In this paper, we present three experiments that address the question of whether the processor initially adopts the analysis that employs the most frequent subcategorisation. To test this, we investigate the way in which people resolve unbounded dependency constructions (such as relative clauses) that contain a local ambiguity about how they should be analysed.

THE USE OF SUBCATEGORISATION PREFERENCES DURING PARSING

Previous experimental evidence about whether subcategorisation preferences affect initial parsing decisions is mixed. As different classes of parsing theories differ in their predictions about whether subcategorisation preferences are used, an answer to this question would be important for

the development of the field. Ferreira and Henderson (1990) had participants read sentences like (1a–b):

1a. He forgot Pam needed a ride home with him.
1b. He wished Pam needed a ride home with him.

The main verb *forgot* preferentially takes a non-phrase object (i.e., the *NP analysis* is preferred), whereas *wished* preferentially takes a sentential complement (i.e., the *SC analysis* is preferred). In both (1a and b), an NP analysis is initially available, as *Pam* can serve as the NP object of the verb. But this analysis becomes impossible at the disambiguating verb *needed*, which is only compatible with the SC analysis. If subcategorisation preferences are initially used, readers should adopt the NP analysis in (1a) but not (1b). However, Ferreira and Henderson found evidence that readers adopted the NP analysis in both cases (essentially, readers were disrupted by both (1a) and (1b) at *needed* to a similar extent as compared with unambiguous control sentences containing the complementiser *that* before *Pam*).

However, Trueswell et al. (1993) found very different results using structurally similar items. In their study, readers experienced difficulty at the disambiguating verb when the main verb was NP preference but not when it was SC preference. It is possible that readers always adopted the NP analysis initially but rapidly reanalysed in the sentences involving NP-preference verbs. Indeed, Trueswell et al. found some increased reading time for such sentences before disambiguation. However, they argued against this explanation, and instead suggested that the SC analysis is more complex than the NP analysis.

Garnsey et al. (1997) identified a problem with Trueswell et al.'s (1993) items: The NP analyses for sentences containing NP-preference verbs were much more plausible than the NP analyses for sentences containing SC-preference verbs. If readers initially adopted the NP analysis but tended to reanalyse when the initial analysis was less plausible (Pickering & Traxler, 1998), then Trueswell et al.'s findings are compatible with late effects of subcategorisation preferences. In fact, Garnsey et al. did not find evidence for difficulty at disambiguation for sentences containing SC-preference verbs, even though they controlled for plausibility. Thus, their results provide some support for early effects of subcategorisation preferences. However, the effects of the plausibility manipulation were less clear, their ambiguous region was short, and their critical finding is based on null result in an experiment with few items and hence low power. (See Pickering, Traxler, & Crocker, 2000, for extensive discussion.)

Pickering et al. (2000) provided evidence that the processor does not base its initial choice of analysis on subcategorisation preferences. In two experiments, participants read sentences like (1):

1a. The young athlete realised her potential one day might make her a world-class sprinter.
1b. The young athlete realised her exercises one day might make her a world-class sprinter.

In (1), the verb *realised* has a preference for a sentential complement (as in (1)) over a noun-phrase object, as in *realised her potential*. However, Pickering et al. found that readers initially adopted the noun-phrase analysis (henceforth, *NP analysis*). They established that (1a) and (1b) are equally plausible on the correct sentential complement analysis, but differ in plausibility on the NP analysis: *The young athlete realised her potential* is plausible, whereas *The young athlete realised her exercises* is implausible (or anomalous). Eye-tracking evidence showed that readers experienced difficulty with (1b) in comparison with (1a) before reading the disambiguating verb *might*. This pattern can only be explained on the assumption that they adopted the NP analysis, despite its relative infrequency. In fact, Kennison (2001) found no evidence for any effect of subcategorisation preferences during the processing of such sentences at all. Sentences that were disambiguated to the NP analysis were easier to process than sentences that were disambiguated to the S analysis, and locally ambiguous S analysis sentences were harder than control sentences disambiguated with *that*, irrespective of whether the verb was NP- or SC-preference.

Mitchell (1987) suggested that readers go one stage further, and ignore subcategorisation information entirely. He had participants read sentences containing verbs like *sneezed* or *visited*:

2a. After the child had sneezed the doctor prescribed a course of injections.
2b. After the child had visited the doctor prescribed a course of injections.

The verb *sneezed* is almost always intransitive except in some rare constructions like *sneezed a big sneeze*. The sentences were presented as two fragments, with the break occurring after *doctor*. Participants took longer to read *After the child had sneezed the doctor* than *After the child had visited the doctor*. However, no difference occurred when an adverbial phrase like *during surgery* appeared before *the doctor*. Mitchell interpreted these results as showing that participants initially treated *the doctor* as the object of *sneezed* (in the absence of an adverbial phrase), but that they then realised that this analysis was impossible and reanalysed. Mitchell also found that the second region, *prescribed a course of injections*, was easier to process in (2a) than (2b), suggesting that participants had already reanalysed in (2a), and that they were surprised by the continuation in (2b)

alone. Adams, Clifton, and Mitchell (1998) suggested that this pattern of results may have been due to the way that the sentences were segmented. Using eye-tracking, they found no evidence of difficulty with sentences like (2a) around *the doctor*. In contrast, they did find difficulty with sentences like (2b) after disambiguation. However, Van Gompel and Pickering (2001) lengthened the ambiguous noun phrase and found results that accorded with Mitchell's findings. If Mitchell's theoretical interpretation of the results is correct, it appears that readers can override very strong subcategorisation preferences and that they have a very strong preference to attach noun phrases as arguments of an available verb if possible (see also Traxler, 2002).

In conclusion, some evidence suggests that readers are guided by subcategorisation frequency in deciding which analysis to adopt (Garnsey et al., 1997; Trueswell et al., 1993). Other evidence, however, suggests that initial decisions are not based on subcategorisation preferences, and that the processor tends to prefer analyses in which a noun phrase is treated as an argument of a verb over analyses in which it does not, even if the former analyses are rarer (Kennison, 2001; Pickering et al., 2000) and perhaps even if they are practically impossible (Mitchell, 1987). In the rest of this paper, we address the role of subcategorisation preferences in relation to one specific type of construction, unbounded dependencies.

SUBCATEGORISATION PREFERENCES AND UNBOUNDED DEPENDENCIES

In an unbounded dependency, two constituents that bear a close linguistic relationship are separated from each other by a considerable amount of intervening material. For example, in (3), *the book* is separated from *about* by five words and, more importantly, a clause boundary, even though the sentence states that the author wrote about a book:

3. We like the book that the author wrote unceasingly about.

Following psycholinguistic terminology, we say that *the book* is the filler associated with the preposition *about* (strictly, this relationship is mediated by the complementiser *that*). In most constructions, verbs and prepositions are very close to their arguments, but in unbounded dependencies, there is no limit to the number of words or clauses that may separate them.

One line of research has asked whether unbounded dependencies are formed in a similar way to other types of dependency or whether their special linguistic status means that they are processed in a special way. Linguists often assume that unbounded dependency constructions are unusually complex. Within transformational grammar, they are a clear example of a construction that involves movement, and therefore have an

underlying structure that differs considerably from the surface structure (Chomsky, 1981). However, there is a good deal of evidence to suggest that they are not particularly complex to process, and in fact are parsed in a comparable way to other dependencies. For example, Traxler and Pickering (1996) found that participants' eye movements were disrupted by an implausible unbounded dependency as soon as they could form the dependency, just as in other eye-tracking experiments concerned with local dependencies (Pickering & Traxler, 1998; Trueswell et al., 1994) or with other linguistic processes such as anaphoric reference (Garrod, Freudenthal, & Boyle, 1994). Further evidence for immediate formation of unbounded dependencies comes from demonstrations of the 'filled-gap' effect (Boland, Tanenhaus, Garnsey, & Carlson, 1995; Bourdages, 1992; Crain & Fodor, 1985; Frazier & Clifton, 1989; Pickering, Barton, & Shillcock, 1994; Stowe, 1986), from cross-modal priming in speech (Nicol, Fodor, & Swinney, 1994; Nicol & Pickering, 1993; Nicol & Swinney, 1989, cf. McKoon, Albritton, & Ratcliff, 1996; McKoon & Ratcliff, 1994; McKoon, Ratcliff, & Ward, 1994), and from studies of word-by-word self-paced reading, where participants at the same time monitored when sentences stopped making sense (Boland et al., 1995; Tanenhaus, Carlson, & Trueswell, 1989, cf. Boland, Tanenhaus, & Garnsey, 1990; Garnsey, Tanenhaus, & Chapman, 1989; Stowe, Tanenhaus, & Carlson, 1991).

Traxler and Pickering (1996) found plausibility effects at the verb by monitoring eye movements during reading. One experiment contrasted the plausible (4a) with the implausible (4b):

4a. That's the pistol with which the heartless killer shot the man yesterday afternoon.
4b. That's the garage with which the heartless killer shot the man yesterday afternoon.

Immediately after they encountered *shot*, participants spent longer reading it in (4b) than (4a). This indicates that people form the dependency before they reach the point at which the filler would occur within an untransformed or canonical sentence (here, after *man*, as in *shot the man with the pistol*). This location is known as the gap (or trace). In accord with the claims of Pickering and Barry (1991; Pickering, 1993, 1994), these results suggest that people do not resolve unbounded dependencies by locating gaps at such points and then associating fillers with gaps (though cf. Gibson & Hickok, 1993). Thus, we can conclude that unbounded dependencies appear to be processed as quickly as other dependencies, and that the mechanisms used to resolve them do not appear to be different from those used to resolve other dependencies.

Let us now ask how the processor resolves unbounded dependencies that contain local ambiguity. Since the processor can resolve unbounded

dependencies rapidly, we would expect it to start the process of ambiguity resolution in unbounded dependencies without delay. In addition, because the mechanisms for processing unbounded dependencies do not appear to be special, the results of investigations into the resolution of ambiguities in unbounded dependencies should be informative about ambiguity resolution in general.

Ambiguity occurs in unbounded dependencies to a similar extent to other constructions. For example, in (1), *the book* could be the (direct) object of *wrote* after *We like the book that the author wrote* ... The NP analysis would turn out to be wrong in this case, so a reader who adopted this analysis would be garden pathed, just as in other types of sentences (e.g., reduced relatives). In such cases, the unbounded dependency is sometimes called 'doubtful', following Fodor (1978).

Experimental evidence clearly demonstrates that the processor sometimes adopts the NP analysis immediately. Traxler and Pickering (1996) had participants read sentences like (5):

5a. We like the book that the author wrote unceasingly and with great dedication about while waiting for a contract.
5b. We like the city that the author wrote unceasingly and with great dedication about while waiting for a contract.

The verb *wrote* can be followed by a noun-phrase object (e.g., *wrote the book*) or a prepositional phrase (e.g., *wrote about the book*). The sentences did not differ in plausibility on the correct prepositional-phrase analysis (henceforth, *PP analysis*). They did differ on the NP analysis, with (3a) being plausible (because writing a book is plausible), and (3b) being implausible (because writing a city is implausible). As soon as participants encountered *wrote*, they experienced difficulty in (3a) versus (3b). This demonstrated that participants formed the NP analysis immediately, and hence that unbounded dependencies can be formed immediately even in cases of local ambiguity. However, there was no evidence that the unbounded dependency was formed if island constraints (Ross, 1967) rendered it ungrammatical (Fodor, 1989; McElree & Griffith, 1998; Pickering et al., 1994; Stowe, 1986, cf. Clifton & Frazier, 1989; Freedman & Forster, 1985), so there do appear to be some restrictions on unbounded dependency formation.

Now consider (6) below:

6. That's the cat that the dog worried about after going to the vet because of an injury.

Like *wrote*, the verb *worried* can be used on the NP analysis. But unlike *wrote*, this analysis is relatively infrequent (see norming below). Our question is whether the processor still adopts the NP analysis immediately,

just as it does in (5). How the processor resolves this ambiguity depends on its strategy in the face of a subcategorisation ambiguity. As in other cases of subcategorisation ambiguity, we can contrast two basic kinds of account: ones where subcategorisation preferences affect initial parsing preferences, and ones where they do not. Here, we specifically focus on two alternatives, which we call the *first-resort* and *lexical-frequency* accounts (see Fodor, 1978). According to the first-resort strategy, the processor immediately adopts the NP analysis. According to the lexical-frequency account, it only adopts this analysis immediately if the verb thereby uses its most frequent subcategorisation frame. Note that this account is deterministic, in the sense that the processor always adopts the most frequent analysis (see General Discussion for consideration of a probabilistic alternative). In (6), the first-resort strategy predicts that the processor immediately treats *the cat* as the object of *worried* (i.e., *the dog worried the cat*); whereas the lexical-frequency account predicts that the processor does not.

With respect to unbounded dependencies, the Garden Path theory, along with most other accounts that are not based on frequency, assumes the first-resort strategy. The processor always forms a potential unbounded dependency at the verb when the verb can take an object argument, however rare the transitive analysis is. Specifically, it proposes the Active filler strategy (Clifton & Frazier, 1989; Frazier & Clifton, 1989; Frazier & Flores D'Arcais, 1989; Mecklinger, Schriefers, Steinhauer, & Friederici, 1995; Schriefers, Friederici, & Kühn, 1995), or the Minimal chain principle (De Vincenzi, 1991). It assumes the existence of traces, and predicts that they are located as early as possible. In contrast, the first-resort strategy without traces does not make predictions that depend on trace location (Pickering, 1993, 1994; Pickering & Barry, 1991; Traxler & Pickering, 1996). These first-resort strategies differ with respect to the processing of unbounded dependencies in verb-final languages and the formation of potential unbounded dependencies in subject position. However, for the ambiguities considered in this paper, the two versions of the first-resort strategy do not differ.[1]

Some frequency-based accounts predict that unbounded dependencies are resolved in accord with the lexical-frequency account. Most specifically, Ford et al. (1982) assumes that initial decisions are based on subcategorisation preferences alone, with the processor adopting the analysis that employs the most frequent subcategorisation frame. Other

[1] Note that we ignore cases where the verb cannot take a (direct) object. It is unclear whether the Garden Path theory predicts that the processor tries to adopt an (impossible) NP analysis in such cases (i.e., whether it ignores subcategorisation information entirely) or whether it merely ignores subcategorisation preferences (see above discussion of Mitchell, 1987).

frequency-based accounts make less clear predictions, because they assume that other information is taken into account. For instance, a particular verb might be predominantly used transitively with one kind of subject (e.g., *the man rolled*) but intransitively with another kind of subject (e.g., *the ball rolled*). Constraint-based accounts (MacDonald et al., 1994; Trueswell et al., 1993) predict that subcategorisation frequency constitutes an important contributor to initial preferences, but that other factors (such as the subject of the verb in the case of *rolled*) may override the impact of subcategorisation frequency. Hence the main concern of this paper is to investigate lexical-frequency accounts.

There is little experimental evidence to distinguish between first-resort and lexical-frequency accounts of the processing of unbounded dependencies. The critical distinction is between verbs that preferentially take a noun-phrase object (henceforth, *NP-preference verbs*) and verbs that do not. Frazier and Clifton (1989) provided some evidence for the first-resort account. Participants read sentences like (7) below:

7a. What did the cautious old man whisper to his fiancée during the movie last night?
7b. What did the cautious old man whisper to his fiancée about during the movie last night?

The verb *whisper* is not NP-preference, so the lexical-frequency strategy predicts that the processor does not form the unbounded dependency. Frazier and Clifton found some evidence that *to his fiancée* in (7a) was easier to read than *to his fiancée about* in (7b), thereby supporting the first-resort account. This may indicate that the processor formed the unbounded dependency in both sentences and subsequently performed reanalysis in (7b). However, there are other explanations of the results, such as the differences between the regions. In addition, it is possible that some of the verbs were actually NP-preference, so the experiment may not adequately test the first-resort strategy.

However, there is also some support for the lexical-frequency account. Stowe et al. (1991) had participants read sentences like (8a–d) below in a word-by-word self-paced reading task:

8a. The teacher wondered which book the students read quietly about.
8b. The teacher wondered which song the students read quietly about.
8c. The teacher wondered which patient the orderly hurried quickly towards.
8d. The teacher wondered which bed the orderly hurried quickly towards.

Participants read the NP-preference verb *read* faster in the plausible sentence (8a) than in the implausible sentence (8b); but they read the verb

hurried, which is not NP-preference, equally fast in (8c) and (8d). These data support lexical frequency. However, the lack of a difference between (8c) and (8d) may have been due to a lack of sensitivity of the technique, to the use of only seven items per condition, or to a potentially weak plausibility manipulation. A first-resort strategy with rapid reanalysis might produce only a small effect with verbs that are not NP-preference. Thus these results suggest that subcategorisation preferences affect some stage of processing, but do not show that they guide initial attachment decisions.

In conclusion, the processor appear to form unbounded dependencies immediately with NP-preference verbs. But it is much less clear what happens with verbs that are not NP-preference. Hence we report two self-paced reading experiments that consider the processing of unbounded dependencies in construction where the filler can act either as the NP object of the verb or can be associated with a preposition following the verb, and employ both NP-preference verbs and verbs that are preferentially followed by a prepositional phrase (henceforth, *PP-preference verbs*). We also report an eye-tracking experiment using PP-preference verbs. The lexical frequency account predicts that the processor forms unbounded dependencies with NP-preference verbs but not PP-preference verbs. In contrast, the first-resort strategy predicts that the processor forms unbounded dependencies immediately for both types of verb.

Our basic logic is to manipulate the plausibility of the NP analysis while keeping the plausibility of the (ultimately correct) PP analysis constant (Pickering & Traxler, 1998; Pickering et al., 2000; Traxler & Pickering, 1996). If people experience difficulty with the sentences that are implausible on the NP analysis (versus sentences that are plausible on this analysis) when they read the critical verb, then they must have computed the NP analysis. If this effect occurs for the sentences containing PP-preference verbs, the lexical-frequency account is disconfirmed. Pickering and Traxler (1998) and Pickering et al. (2000) (in two out of three experiments) also found that difficulty reversed after disambiguation, with sentences with plausible initial analyses being harder than sentences with implausible initial analyses. Pickering and Traxler argued that people commit more strongly to plausible than implausible analyses, so when they find they have to reanalyse, they experience more difficulty abandoning a plausible than an implausible (mis)analysis. However, Traxler and Pickering only found weaker evidence for reversal of difficulty after disambiguation. Thus, if people adopt the NP analysis immediately, a secondary prediction is that they should experience more difficulty after disambiguation for the plausible than the implausible sentences.

EXPERIMENT 1

Method

Participants. Thirty-eight undergraduates at The Florida State University participated for course credit. All were native American English speakers and had normal or corrected-to-normal vision.

Stimuli. Participants read 28 sentences like (9a and b) and (10a and b):

9a. That's the cat that the dog worried compulsively about after going to the vet because of an injury. (PP-preference, plausible)
9b. That's the car that the dog worried compulsively about after going to the vet because of an injury. (PP-preference, implausible)

10a. That's the general that the soldier killed enthusiastically for during the war in Korea. (NP-preference, plausible)
10b. That's the country that the soldier killed enthusiastically for during the war in Korea. (NP-preference, implausible)

In (9a), the NP analysis is plausible at *worried*, because a dog can worry a cat; in (9b), it is implausible, because a dog cannot worry a car. But in both (9a) and (9b), the PP analysis is plausible, because a dog can worry about a cat or about a car. Similarly, in (10a), the NP analysis is plausible at *killed* because a soldier can kill a general; in (10b), it is implausible, because a soldier can not kill a country. But in both (10a) and (10b), the PP analysis is plausible, because a soldier can kill for a general or for a country (see plausibility norming below). Notice also that the NP analysis becomes impossible following *after going to the vet* in (9) and after *for during the war* in (10).

The filler nouns (e.g., *cat/car*, *general/country*) were matched for length and frequency across verb type and plausibility of the NP analysis on the Kučera and Francis (1967) word frequency norms. A 2 Plausibility (plausible vs. implausible NP analysis) $\times$ 2 Verb Preference (PP- vs. NP-preference) ANOVA treating both factors as within items produced no main effects and no interaction (all $p > .15$). Likewise, no simple main effects attained statistical significance (all $p > .15$).

We conducted six norming studies, all of which employed native speakers of American English. The first three are concerned with making sure that plausibility varies appropriately, with the plausibility of the NP analysis differing between conditions, but the plausibility of the PP analysis remaining relatively constant. The last three are concerned with determining the preferences for the two relevant analyses. For the lexical frequency account, the critical data are the verb norming data, where subcategorisation preferences are assessed. The final two studies con-

sidered the preferences for NP and PP analyses given either the verb and its subject, or the whole prior context, including the manipulation of the filler. These data are relevant to frequency-based accounts in general.

Plausibility norming. To determine the plausibility of the NP analyses, 20 further participants rated the plausibility of 72 sentences like (11a and b) on a scale from 0 (makes no sense) to 7 (makes perfect sense).

11a. That's the cat that the dog worried compulsively.
11b. That's the car that the dog worried compulsively.

We eliminated materials whenever the mean rating of the plausible sentence (e.g., 11a) was below 5 or the mean rating of the implausible sentence (e.g., 11b) was above 2. For the NP-preference items, the mean rating of the plausible sentences was 5.5, and the mean rating of the implausible sentences was 1.4; for the PP-preference items, the mean rating of the plausible sentences was 5.7, and the mean rating of the implausible sentences was 1.2.

To determine the plausibility of the PP analyses, the same group rated the plausibility of sentences like (12a and b):

12a. The dog worried compulsively about the cat.
12b. The dog worried compulsively about the car.

The plausibility of the correct analysis did not differ across conditions. For sentences with plausible NP analyses, the mean rating was 5.0; for sentences with implausible NP analyses, the mean rating was 5.3 (both $p > .20$).

After selecting the materials for the experiment, we conducted a further plausibility test on sentences like (13):

13a. The dog worried the cat compulsively.
13b. The dog worried the car compulsively.
13c. The dog worried about the cat compulsively.
13d. The dog worried about the car compulsively.

We constructed four lists of sentences, so that each participant saw exactly one version of each item and equal numbers of items in all conditions. Thirty-nine participants provided ratings under the same instructions as the previous plausibility test. The differences between this plausibility post-test and the plausibility pre-test are that none of the test sentences contain unbounded dependencies, as it is conceivable that the process of forming an unbounded dependency might influence raters' plausibility judgements (compare (13a and 13b) to (11a and 11b) above), and that the adverb now occurs after the object noun phrase. Plausibility post-test data confirmed the pattern observed in the plausibility pre-test: There was a

large difference in plausibility of the sentences derived from the NP analysis, with the mean rating being 5.4 for sentences like (13a) with plausible NP analyses versus 1.9 for sentences like (13b) with implausible NP analyses. However, there was no difference for the sentences derived from the PP analysis, with the mean rating being 4.9 for sentences like (13c) (associated with sentences with plausible NP analyses) versus 4.4 for sentences like (13d) (associated with sentences with implausible NP analyses).

Verb norming. To determine the subcategorisation preferences of the verbs, we presented 20 participants with a total of 44 verbs in a fixed random order. Writers read and followed these instructions:

> Here is a list of words. Write one sentence containing each word. Please use the word exactly as it appears in the list. For example, if the word is ''stopped'', then that exact word should appear somewhere in the sentence you write. Please try to write complete, grammatical sentences that make sense.

A second set of 70 writers saw a different set of 65 verbs. Table 1 presents the results of these two norming sessions for the verbs employed in the items that were presented during the self-paced reading phase of the experiment, and for the other norming studies reported below. We computed the proportion of PP responses by dividing the number of PP responses by the sum of the PP and NP responses (other kinds of responses are also presented in Table 1 for completeness). The PP-preference verbs had PP proportions between 0.6 and 1.0, with a mean proportion of 0.78. The NP-preference verbs had PP proportions between 0.0 and 0.44, with a mean proportion of 0.12.

Subject-plus-verb norming. Twenty further participants completed fragments like *The dog worried* ... in a random order (along with 28 filler sentences which intervened between each experimental item). The PP-preference fragments had PP proportions between .16 and .95, with a mean proportion of .61. The NP-preference fragments had PP proportions between .00 and .94, with a mean proportion of .21. This norming task suggests that the structural preferences that readers have for subject plus verb may be different from those for the verb on its own. Most importantly, some PP-preference verbs may show a different preference when combined with the subject noun phrases that we employed. This difference is not critical for the lexical-frequency account as defined above (following Ford et al., 1982, and discussion in Fodor, 1978), but obviously might be relevant to an alternative account in which the impact of the

TABLE 1
Norming results

	Verb norming results			*Subject-plus-verb norming results*		
	Sentence type (%)			*Sentence type (%)*		
	NP	*PP*	*Other*	*NP*	*PP*	*Other*
PP-preference verbs						
landed	25	40	35	80	15	5
blabbed	30	45	25	25	30	45
communicated	10	50	40	35	45	20
fished	5	45	50	20	55	25
pointed	10	90	0	5	95	0
preached	30	45	25	25	60	15
raced	35	55	10	45	35	20
searched	15	75	10	70	15	15
shouted	10	50	40	25	30	45
spoke	0	55	45	5	35	60
travelled	20	40	40	30	70	0
worried	0	50	50	15	70	15
argued*	11	64	24	45	35	20
swore*	6	17	77	5	80	15
NP-preference verbs						
killed	70	0	30	95	5	0
asked	85	10	5	25	15	60
called	65	10	25	95	0	5
heard	75	10	15	90	10	0
hid	50	40	10	30	65	5
instructed	65	0	35	90	5	5
kidnapped	45	0	55	90	5	5
paid	50	10	40	70	30	0
operated*	50	34	16	5	75	20
read	80	5	15	95	5	0
taught	90	0	10	80	5	15
threw	80	5	15	90	0	10
told	85	5	10	75	25	0
wrote	75	15	10	95	5	0

Note: * indicates that these verbs were included in the second norming session (with 70 participants). *NP* means that an NP object appeared immediately following the verb in the writer's continuation, as in *Sally blabbed the secret*. *PP* means that a PP appeared immediately after the verb, as in *Sally blabbed about the secret*. *Other* indicates that the response did not fit into one of the other categories (and includes illegible and ungrammatical responses).

subject on the verb's subcategorisation preferences is taken into account (cf. Hare, McRae, & Elman, 2003).

Whole fragment completions. Forty further participants completed the actual fragments used in the experiment up to the critical verb (e.g., *That's*

the cat/car that the dog worried, see Table 2). Participants saw exactly one version of each fragment (plausible or implausible, with respect to the NP analysis), and equal numbers of items in all conditions. The mean PP proportions were: .35 for the plausible PP-preference fragments, .90 for the implausible PP-preference fragments, .12 for the plausible NP-preference

TABLE 2
Whole fragment norming results

	When Filler was Plausible Direct Object			*When Filler was Implausible Direct Object*		
	Sentence type (%)			*Sentence type (%)*		
	NP	*PP*	*Other*	*NP*	*PP*	*Other*
PP-preference verbs						
landed	100	0	0	15	85	0
blabbed	95	5	0	25	75	0
communicated	80	20	0	5	95	0
fished	5	95	0	65	35	0
pointed	75	25	0	0	90	10
preached	75	25	0	5	90	5
raced	90	10	0	25	70	5
searched	55	35	10	10	90	0
shouted	100	0	0	0	95	5
spoke	100	0	0	0	100	0
travelled	35	65	0	0	90	10
worried	10	90	0	0	100	0
argued*	70	25	5	0	100	0
swore*	20	80	0	0	100	0
NP-preference verbs						
killed	100	0	0	10	50	40
asked	80	15	5	15	80	5
called	100	0	0	85	15	0
heard	95	5	0	0	70	30
hid	80	20	0	5	90	5
instructed	70	15	15	20	80	0
kidnapped	100	0	0	25	45	30
paid	85	15	0	20	80	0
operated*	30	65	5	0	95	5
read	100	0	0	0	100	0
taught	100	0	0	10	90	0
threw	100	0	0	5	95	0
told	95	5	0	0	95	5
wrote	100	0	0	0	100	0

Note: *NP* means that the participant treated the filler as the object of the main verb (as in, e.g., *That's the plane that the pilot landed at the airport*). *PP* means that the participant treated the filler as part of a prepositional-phrase or infinitival complement to the main verb (as in, e.g., *That's the truck that the pilot landed the plane behind*).

fragments, and .86 for the implausible NP-preference fragments. A 2 Verb Preference (PP- vs. NP-preference) × 2 Plausibility (plausible vs. implausible NP analysis) ANOVA treating items as a random factor revealed main effects of verb preference, $F_2(1, 26) = 4.39$, $p < .05$, MSE = 22.3, and plausibility, $F_2(1, 26) = 91.6$, $p < .0001$, MSE = 25.3, but no interaction between these factors. Clearly, completion preferences are changed quite dramatically by the inclusion of the whole fragment, and differ depending on the plausibility of the NP analysis (even though the plausibility of the PP analysis does not change).

Procedure. One version of each test sentence was randomly assigned to one of two lists, so that each list contained 28 items. Fourteen of the items contained PP-preference verbs, and fourteen contained NP-preference verbs. Half of the sentences in each group had plausible object analyses; and half had implausible object analyses. The test sentences were presented along with 90 other sentences of varying types. At least one non-experimental sentence intervened between each experimental item. Slash marks in sentence (9a) (repeated here) indicate region boundaries. (The slash marks did not appear in the items presented to the readers.)

9a. That's the cat/ that the dog/ worried compulsively/ about after going to the vet/ because of an injury.

We chose this segmentation because we needed to measure effects in the ambiguous region separately from the disambiguating region. Inclusion of adverbs like *compulsively* allowed us to capture spillover effects caused by the preceding verb, and thus maximised our chances of detecting differences between conditions before disambiguation. Such differences are critical to the lexical-frequency account (but note that exclusion of the adverbial did not alter the pattern of results, see Experiment 2 below). This disambiguating region always included a phrase like *after going to the vet* that was sometimes required to eliminate the NP analysis. Without this material, some of the sentences could continue as in, for example, *That's the cat that the dog worried compulsively about ten times a day*, which could mean that the cat worried the dog compulsively about ten times a day (i.e., treating *the cat* as the NP object of *worried*). Half of the participants were randomly assigned to each list. The test sentences were presented on a MacIntosh PC running PsyScope presentation software (Cohen, MacWhinney, Flatt & Provost, 1993). Participants were instructed to read at their normal pace and comprehend the sentences to the best of their ability. Each trial began with a 'Ready' signal on the screen. Participants pressed the space bar on the keyboard to view the first segment of the sentence, which was presented in the centre of the screen. When they

finished reading the first segment, they pressed the space bar and the next segment replaced the first segment. After reading all of the segments, a yes-or-no comprehension question (e.g., *Did cat worry the dog?*) replaced the last segment of the sentence in the middle of the screen. Participants responded to the question by pressing a key labelled 'yes' or a key labelled 'no'. The computer recorded the time between successive key presses and responses to the comprehension questions.

Results and discussion

Table 3 presents mean reading time by region and condition for Experiment 1, along with the proportion of comprehension questions answered correctly. No main effects or interactions occurred in the comprehension data (all Fs < 1), so we will not consider them further. Before analysing the reading time data, we removed data that fell more than two standard deviations above or below the condition means. This procedure eliminated 5.0% of the responses.

If the lexical-frequency account is correct, then participants would not attempt to form the NP analysis for PP-preference verbs. Thus, plausibility of the NP analysis should not affect processing of sentences with PP-preference verbs (because the NP analysis will never be attempted). By

TABLE 3
Experiment 1: Mean reading time by region and condition (means after removing four items appear in parentheses)

	Region					
Verb preference	*1*	*2*	*3*	*4*	*5*	*Proportion correct*
PP-preference plausible	984 (972)	890 (878)	1012 (1045)	1309 (1392)	971 (893)	0.87
PP-preference implausible	945 (880)	893 (948)	1113 (1157)	1228 (1326)	873 (867)	0.88
NP-preference plausible	853	958	1045	1392	883	0.85
NP-preference implausible	880	948	1157	1326	869	0.86

Example sentences (with slashes indicating region breaks):

PP-preference plausible: That's the plane/that the pilot/landed carefully/behind in the fog/at the airport.

PP-preference implausible: That's the truck/that the pilot/landed carefully/behind in the fog/at the airport.

NP-preference plausible: That's the general/that the soldier/killed enthusiastically/for during the war/in Korea.

NP-preference implausible: That's the country/that the soldier/killed enthusiastically/for during the war/in Korea.

contrast, participants would attempt to adopt the NP analysis for NP-preference verbs. This should produce longer reading times in the ambiguous region (region 3) when the incorrect reading is implausible than when the incorrect reading is plausible; this pattern should reverse in the disambiguating region (region 4): see Pickering and Traxler (1998; Pickering et al., 2000). Thus, we first subjected the data to 2 Verb Preference (PP- vs. NP-preference) × 2 Plausibility (plausible vs. implausible NP analysis) × 2 Region (region 3 vs. region 4) ANOVAs. All factors were within participants, verb preference was manipulated between items, and the other factors were within items. If readers follow the lexical-frequency account, we should observe a three-way interaction of verb preference, plausibility, and region. In fact, there was not the least hint of this interaction, $F_1(1, 37) = 0.002$, ns, MSE = 26030; $F_2(1, 26) = 0.328$, ns, MSE = 10600. There was, however, a powerful interaction of plausibility and region, $F_1(1, 37) = 17.3$, $p < .001$, MSE = 35380; $F_2(1, 26) = 19.85$, $p < .001$, MSE = 10600. Sub-analyses indicated that plausibility and region interacted in sentences with PP-preference verbs, $F_1(1, 37) = 8.19$, $p < .01$, MSE = 38062; $F_2(1, 13) = 11.05$, $p < .01$, MSE = 12128, and in sentences with NP-preference verbs, $F_1(1, 37) = 12.86$, $p < .001$, MSE = 23348; $F_2(1, 13) = 8.81$, $p < .01$, MSE = 9072.

Tests for simple effects revealed the source of these interactions of Plausibility and Region. For PP-preference verbs, sentences with implausible object analyses produced longer reading times in the ambiguous region than sentences with plausible object analyses, $F_1(1, 37) = 5.06$, $p < .04$; $F_2(1, 13) = 6.57$, $p < .03$. There is a strong suggestion that this effect reversed in the disambiguating region, $F_1(1, 37) = 3.23$, $p < .09$; $F_2(1, 13) = 4.57$, $p < .05$. Sentences with NP-preference verbs produced a similar main effect of plausibility in the ambiguous region, $F_1(1, 37) = 10.16$, $p < .01$; $F_2(1, 13) = 9.64$, $p < .01$, but the reversal in the disambiguating region was marginal by subjects alone, $F_1(1, 37) = 3.55$, $p < .07$, $F_2(1, 13) = 1.19$, ns.

Additionally, we analysed the data from Regions 3 and 4 separately, using a 2 Verb Preference × 2 Plausibility design. Region 3 produced an effect of plausibility, $F_1(1, 37) = 31.5$, $p < .0001$; $F_2(1, 26) = 10.3$, $p < .01$, but no hint of an interaction, $F_1(1, 37) = 0.07$, ns; $F_2(1, 26) = 0.01$, ns. Likewise, Region 4 produced an effect of Plausibility, $F_1(1, 37) = 4.11$, $p < .05$; $F_2(1, 26) = 4.11$, and no interaction, $F_1(1, 37) = 0.06$, ns; $F_2(1, 26) = 0.62$, ns. In short, sentences with PP-preference verbs produced plausibility effects every bit as large as sentences with NP-preference verbs.

It could be argued that the preceding results were skewed by having four items included in the PP-preference category that produced greater NP-object proportions in the fragment norming. As we have pointed out, this

does not affect the predictions of the lexical-frequency account, but our results would help rule out a broader class of account if they were not affected by differences between verb norming and subject-plus-verb norming. To see whether this was the case, we re-analysed the data after removing the four items from the PP-preference category that were classified as NP-preference in the stem-completion task. Eliminating the four items did not change the results in any of the analyses. There was still no three-way interaction, $F_1(1, 37) = 0.055$, ns, MSE = 24195; $F_2(1, 22) = 0.130$, ns, MSE = 9771, but there was still an interaction of plausibility and region, $F_1(1, 37) = 18.3$, $p < .0001$, MSE = 29768; $F_2(1, 21) = 16.4$, $p < .001$, MSE = 9771. When the remaining 10 PP-preference items were analysed separately, they also produced an interaction of plausibility and region, $F_1(1, 37) = 12.9$, $p < .001$, MSE = 23348; $F_2(1, 9) = 7.57$, $p = .02$, MSE = 10780. This interaction occurred because PP-preference items with implausible misanalyses produced longer reading times in the ambiguous region than PP-preference items with plausible misanalyses, 1045 versus 1157 ms: $F_1(1, 37) = 10.2$, $p < .01$; $F_2(1, 9) = 10.7$, $p < .01$. The subjects analysis of the disambiguating region produced a trend toward a reversal of this effect, 1392 versus 1326 ms: $F_1(1, 37) = 3.55$, $p = .07$, which was not significant in the items analysis, $F_2(1, 9) < 1$. These analyses demonstrate that the theoretically critical results were not due to effects occurring in the four PP-preference items that were NP preference in the fragment completion task. Additionally, we conducted 2 Plausibility $\times$ 2 Verb Preference analyses for Regions 3 and 4 separately. Region 3 produced a main effect of plausibility, $F_1(1, 37) = 31.1$, $p < .0001$; $F_2(1, 22) = 11.4$, $p < .01$, but no interaction, $F_1(1, 37) = 0.63$, ns; $F_2(1, 22) = 0.27$, ns. Region 4 produced no main effects or interactions in the reduced set of items.

Plausibility effects in Region 3 for both types of verbs show that readers adopted the NP analysis, regardless of the lexical characteristics of the verb. These results, therefore, directly challenge the lexical-frequency account, in which the processor uses subcategorisation preferences to guide its attachment decisions in sentences with unbounded dependencies.

EXPERIMENT 2

One concern with Experiment 1 is that its results conflict fairly directly with Stowe et al. (1991). Specifically, we found plausibility effects for PP preference verbs, but they did not; and they found an interaction between verb preference and plausibility, but we did not. To determine whether we are correct, it is necessary to replicate our findings. A second issue with Experiment 1 is that it is conceivable that readers did not form the unbounded dependency at the verb, but waited until reading the adverb following the verb (e.g., *carefully*) before forming this association. The

strongest prediction of the first-resort strategy is that the plausibility effect would occur on the verb itself. In Experiment 2, therefore, the sentences appeared without the adverb.

Method

Participants. Forty further undergraduates at The Florida State University participated for course credit.

Stimuli. Participants read 28 sentences like (14a and b) and (15a and b):

14a. That's the cat that the dog worried about after going to the vet because of an injury.
14b. That's the car that the dog worried about after going to the vet because of an injury.

15a. That's the general that the soldier killed for during the war in Korea.
15b. That's the country that the soldier killed for during the war in Korea.

The sentences were identical to those used in Experiment 1, except that the adverb following the critical verb (e.g., *worried* or *killed*) was removed (see Appendix).

Procedure. The procedure was identical to Experiment 1, except that the sentences were presented in four fragments, and the critical verb was included in the second fragment (see 14a):

That's the cat/ that the dog worried/ about after going to the vet/ because of an injury. (14a)

This segmentation was chosen for the same reasons as in the previous experiment.

Results and discussion

Table 4 presents mean reading time by region and condition for Experiment 2, along with the proportion of comprehension questions answered correctly. No main effects or interactions occurred in the comprehension data (all $F < 2.14$), so we will not consider them further. Before analysing the reading time data, we removed data that fell more than two standard deviations above or below the condition means. This procedure eliminated 4.4% of the responses.

TABLE 4
Experiment 2: Mean reading time by region and condition

	Region				
Verb preference	*1*	*2*	*3*	*4*	*Proportion correct*
PP-preference plausible	972 (1017)	1166 (1198)	1435 (1413)	961 (1016)	0.91
PP-preference implausible	966 (1004)	1285 (1298)	1331 (1324)	946 (980)	0.90
NP-preference plausible	827	1206	1413	911	0.88
NP-preference implausible	872	1227	1344	868	0.90

Example sentences (with slashes indicating region breaks):

PP-preference plausible: That's the plane/that the pilot landed/behind in the fog/at the airport.

PP-preference implausible: That's the truck/that the pilot landed/behind in the fog/at the airport.

NP-preference plausible: That's the general/that the soldier killed/for during the war/in Korea.

NP-preference implausible: That's the country/that the soldier killed/for during the war/in Korea.

Table 4 indicates that the results for Experiment 2 were broadly similar to Experiment 1. We subjected the data to 2 Verb Preference (PP- vs. NP-preference) $\times$ 2 Plausibility (plausible vs. implausible NP analysis) $\times$ 2 Region (region 2 vs. region 3) ANOVAs. As in the previous experiment, there was no three-way interaction between verb preference, plausibility, and region, $F_1(1, 39) = 1.99$, ns, MSE = 43600; $F_2(1, 26) = 2.42$, ns, MSE = 16173. (Note that this trend is in the opposite direction to that predicted by lexical frequency.) However, there was a clear interaction of plausibility and region, which corresponds to the familiar cross-over pattern, $F_1(1, 39) = 21.25$, $p < .001$, MSE = 22947; $F_2(1, 26) = 9.90$, $p < .01$, MSE = 16173. As in Experiment 1, we found the interaction between plausibility and region for PP-preference verbs, $F_1(1, 39) = 18.01$, $p < .001$, MSE = 27353; $F_2(1, 13) = 9.68$, $p < .01$, MSE = 16173. For those items containing PP-preference verbs, sentences with implausible object analyses produced longer reading times in the ambiguous region than sentences with plausible object analyses, $F_1(1, 39) = 10.27$, $p < .01$, MSE = 27353; $F_2(1, 13) = 6.04$, $p < .03$, MSE = 16173; and this effect reversed in the disambiguating region, $F_1(1, 39) = 7.82$, $p < .01$, MSE = 27353; $F_2(1, 13) = 3.78$, $p < .07$, MSE = 16173.

Evidence for the formulation of unbounded dependencies was weaker for NP-preference verbs. Although the pattern of means was the same for both PP- and NP-preference verbs, the sentences containing NP-

preference verbs did not produce an interaction of plausibility and region, $F_1(1, 39) = 2.08$, ns, MSE = 39194; $F_2(1, 13) = 1.48$, ns, MSE = 13895. Likewise, sentences with NP-preference verbs did not produce plausibility effects in region 2 (both Fs < 1) or region 3, $F_1(1, 39) = 2.43$, $p < .13$, MSE = 39194; $F_2(1, 13) = 1.24$, $p > .28$, MSE = 13895.

Hence, the cross-over pattern is reliable for the PP-preference verbs but not the NP-preference verbs in this experiment. However, there is no three-way interaction (see above), and no interactions between preference and plausibility on either region: region 2, $F_1(1, 39) = 3.24$, ns, MSE = 29071; $F_2(1, 26) = 1.60$, ns, MSE = 21619; region 3, $F_1(1, 39) = 0.25$, ns, MSE = 48559; $F_2(1, 26) = 0.223$, ns, MSE = 39575). There is no particular theoretical reason to expect a clearer effect with PP- than NP-preference verbs, and Experiment 1 did show the cross-over pattern for both types of verb. More importantly, this experiment replicated the important finding from Experiment 1, which demonstrated that participants adopted the NP analysis for the PP-preference verbs.

As in the previous experiment, we analysed the data from the two critical regions separately using a 2 Plausibility $\times$ 2 Verb Preference design. Region 2 produced a main effect of plausibility, $F_1(1, 39) = 10.8$, $p < .01$; $F_2(1, 26) = 3.80$, $p = 0.6$. The interaction, while marginally significant in the by-subjects analysis, $F_1(1, 39) = 3.24$, $p = .08$, was not significant in the by-items analysis, $F_2(1, 26) = 1.60$, ns. Region 3 produced a main effect of plausibility in the by-subjects analysis, $F_1(1, 39) = 6.89$, $p = .01$, but not in the by-items analysis, $F_2(1, 26) = 1.97$, ns. The interaction was not significant, $F_1(1, 39) = 0.25$, ns; $F_2(1, 26) = 0.22$, ns.

Also as in the previous experiment, the data were reanalysed after removing the four items whose preferences changed from PP-preference to NP-preference in the stem-completion task. Again, there was no three-way interaction, $F_1(1, 39) = 1.22$, ns, MSE = 39439; $F_2(1, 22) = 1.30$, ns, MSE = 16136, but there was an interaction of region and plausibility, $F_1(1, 39) = 12.75$, $p < .001$, MSE = 30511; $F_2(1, 22) = 5.94$, $p = .02$, MSE = 16136, which also occurred when the ten remaining PP-preference items were analysed by themselves, $F_1(1, 39) = 8.07$, $p < .01$, MSE = 22879; $F_2(1, 9) = 4.97$, $p < .05$, MSE = 19373. In the ambiguous region, the ten remaining PP-preference items produced an effect of plausibility in the by-participants analysis: plausible misanalysis, 1198 ms; implausible misanalysis, 1298 ms; $F_1(1, 39) = 8.74$, $p < .01$, and a trend towards an effect of plausibility in the by-items analysis, $F_2(1, 9) = 3.58$, $p = .09$. The disambiguating region did not produce an effect of plausibility in the ten remaining PP-preference items, $F_1(1, 39) = 1.13$, ns, MSE = 22878; $F_2(1, 9) = 1.59$, ns, MSE = 19373 (although the means were in the appropriate direction, with the plausible condition at 1016 ms and the implausible condition at 980 ms). When Regions 2 and 3 were analysed separately in a

2 Plausibility × 2 Verb Preference design, Region 2 produced only a main effect of plausibility, $F_1(1, 39) = 5.43$, $p < .05$; $F_2(1, 26) = 3.26$, $p = .08$. The interaction was not significant, $F_1(1, 39) = 1.85$, ns; $F_2(1, 22) = 1.56$, ns. Region 3 produced an effect of Plausibility only in the by-subjects analysis, $F_1(1, 39) = 4.78$, $p < .05$; $F_2(1, 22) = 1.01$, ns. The interaction was not significant, $F_1(1, 39) = .010$, ns; $F_2(1, 22) = 0.05$, ns. Hence, the effects in the reduced set of PP-preference items were similar to the complete set with respect to the critical effects in the ambiguous region.

EXPERIMENT 3

It is just possible that the immediate effects of plausibility that occurred in Experiments 1 and 2 were somehow the result of specific reading strategies associated with the phrase-by-phrase self-paced-reading task. For example, participants might form an unbounded dependency when the point at which the dependency could be formed corresponded to a region break. To rule out this possibility, we conducted an eye-tracking experiment using the PP-preference items alone (as these constitute the critical conditions to distinguish the accounts). In this respect, the experiment is similar to those reported in Pickering et al. (2000), where the analysis whose plausibility was manipulated was both the less frequent analysis and ultimately not the correct analysis. If the immediate effects of plausibility still emerge under these conditions, it would provide further evidence against the lexical-frequency account.

Method

Participants. Twenty-two undergraduates at the University of South Carolina participated in return for credit toward a course requirement. All of the participants were native speakers of American English and had normal vision and hearing.

Items. The items were the PP-preference items from Experiment 1 (see (9)). The items were randomised and assigned to one of two lists for presentation to readers such that an equal number of each condition appeared in each list and such that no participant was exposed to more than one version of each item. Eighty-seven filler items of various syntactic types were also presented (including seven NP-preference items from Experiment 1).

Procedure. A Fourward Technologies Dual-Purkinje Eye-Tracker monitored participants' eye movements while they read the test sentences. The tracker has angular resolution of 10′ of arc. The tracker monitored only the right eye's gaze location. A PC displayed materials on a VDU 70 cm from participants' eyes. The location of participants' gaze location was

sampled every millisecond and the PC software recorded the tracker's output to establish the sequence of eye fixations and their start and finish times. At the beginning of the experiment, the experimenter seated the participant at the eye tracker and used a bite plate and head rests to minimise head movements and then aligned and calibrated the eye tracker. After reading each sentence, the participant pressed a key. After some of the filler sentences, the participant responded to a comprehension question. Participants received feedback on their responses. All of the participants in the analyses reported below scored at 90% accuracy or above on the comprehension questions. Between each trial, a pattern of boxes appeared on the computer screen along with a cursor that indicated the participants' current gaze location. If the tracker was out of alignment, the experimenter recalibrated it before proceeding with the next trial.

Results and discussion

We report four standard dependent measures: first-pass time, regression-path time, first-pass regressions, and total time. *First-pass time* is the sum of all the fixations beginning with the reader's first fixation in a region until the reader's gaze leaves the region (on single-word regions, first-pass time is equivalent to the gaze duration measure; Rayner & Duffy, 1986). *Regression-path time* is the sum of all the fixations starting with the first fixation in a region until the reader fixates anything to the right of the region, including fixations following a first-pass regression (Brysbaert & Mitchell, 1996; Traxler, Bybee, & Pickering, 1997). A *first-pass regression* occurs when the reader's gaze crosses the left edge of the scoring region following a first-pass fixation. *Total time* is the sum of all of the fixations within a region. Two critical regions were examined. The *ambiguous* region included the critical verb and the following adverb (e.g., *landed carefully*). The *disambiguating region* included the preposition and the following constituent (e.g., *behind in the fog*). In addition, data are reported from the beginning of the sentence (the entire sentence up to the beginning of the ambiguous region) and the end of the sentence (all of the sentence following the disambiguating region). No effects occurred in these two regions, and so we will not discuss them further. All scoring regions included the character space immediately to the left of the first word in the region. Table 5 presents mean scores on four dependent measures by condition for the two scoring regions.

Prior to analysis, all fixation times of 0 ms or greater than 2000 ms were eliminated (4% of the data). The remaining data were subjected to a series of ANOVAs testing for an effect of plausibility and region, with both factors treated as within-participants and within-items. The first-pass data produced a trend toward an interaction of plausibility and region in the by-

TABLE 5
Experiment 3: Mean values on four dependent measures by condition

Example sentence: *That's the plane/truck that the pilot landed carefully behind in the fog at the airport.* (The word preceding the '/' appeared in the plausible misanalysis version of the sentence; the word following the '/' appeared in the implausible misanalysis version.)

First-pass time (ms.)	Beginning	Ambiguous Region	Disambiguating Region	End
Plausible misanalysis	1328	604	703	445
Implausible misanalysis	1287	684	714	432
Regression-path-time (ms)				
Plausible misanalysis	1328	**643**	796	632
Implausible misanalysis	1287	**743**	799	592
First-pass regressions (%)				
Plausible misanalysis	–	4.5	9.7	14.9
Implausible misanalysis	–	7.1	12.3	15.6
Total time (ms)				
Plausible misanalysis	1375	**679**	837	554
Implausible misanalysis	1362	**810**	824	571

Example sentences (with slashes indicating analysis regions):
Plausible: That's the plane that the pilot/landed carefully/behind in the fog/at the airport.
Implausible: That's the truck that the pilot/landed carefully/behind in the fog/at the airport.
Note: Bold type indicates difference significant by both F_1 and F_2.

items analysis, $F_2(1, 13) = 3.62$, $p = .08$, MSE = 4865, but not in the by-participants analysis, $F_1(1, 21) = 2.72$, $p > .10$, MSE = 9042. There was an effect of plausibility in the ambiguous region, as sentences with implausible misanalyses produced longer first-pass times than sentences with plausible misanalyses, $F_1(1, 21) = 7.53$, $p < .01$; $F_2(1, 13) = 7.83$, $p = .02$. The disambiguating region did not produce an effect of plausibility (Fs < 1). The regression-path time analyses produced a trend toward an interaction of plausibility and region in the by-participants analysis alone, $F_1(1, 21) = 3.20$, $p = .09$, MSE = 16278; $F_2(1, 13) = 2.51$, ns, MSE = 8421. Tests for simple effects showed that there was an effect of plausibility in the ambiguous region, $F_1(1, 21) = 6.77$, $p < .02$; $F_2(1, 13) = 6.77$, $p < .02$, but there were no effects in the disambiguating region (Fs < 1). The total time data produced an interaction of plausibility and region, $F_1(1, 21) = 10.2$, $p < .01$, MSE = 11160; $F_2(1, 13) = 13.6$, $p < .01$, MSE = 4555. There was a robust effect of plausibility in the ambiguous region, $F_1(1, 21) = 16.9$, $p < .001$; $F_2(1, 13) = 21.9$, $p < .001$, but again no effects in the disambiguating region (Fs < 1). There were no interactions or main effects in the first-pass regression data (all Fs < 1.20).

As in the previous two experiments, we reanalysed all of the data after removing the four items that were classified as PP-preference based on the

verb norming task but NP-preference based on the stem completion task. The first-pass data from the reduced set did not produce an interaction of plausibility and region, but did produce a strong trend toward an effect of plausibility in the ambiguous region: plausible misanalysis, 634 ms; implausible misanalysis, 704 ms; $F_1(1, 21) = 3.97$, $p = .06$, MSE = 13426; $F_2(1, 9) = 3.69$, $p = .09$, MSE = 2904. No effects occurred in the first-pass data from the disambiguating region: plausible misanalysis, 676 ms; implausible misanalysis, 733 ms; $F_1(1, 21) = 2.67$, $p = .12$, MSE = 13426; $F_2(1, 9) = 1.60$, ns, MSE = 2904. Similarly, the regression-path time data did not produce an interaction of plausibility and region (all Fs < 1.22). There was a trend toward a plausibility effect in the regression-path-time data from the ambiguous region in the by-participants analysis alone: plausible misanalysis, 676 ms; implausible misanalysis, 777 ms; $F_1(1, 21) = 3.58$, $p = .07$, MSE = 31193; $F_1(1, 9) = 2.40$, $p = .16$, MSE = 10837. The regression-path-time data did not produce an effect of plausibility in the disambiguating region (Fs < 1). The total time data from the reduced set of items produced an interaction of plausibility and region, $F_1(1, 21) = 4.41$, $p < .05$, MSE = 20986; $F_2(1, 9) = 6.98$, $p = .03$, MSE = 4462. Sentences with implausible misanalyses produced longer total time in the ambiguous region than sentences with plausible misanalyses, 855 ms vs. 727 ms; $F_1(1, 21) = 8.58$, $p < .01$; $F_2(1, 9) = 14.3$, $p < .01$, but no effects occurred in the disambiguating region (both Fs < 1, ns, plausible misanalysis = 851 ms, implausible misanalysis = 850 ms). The results from the reduced set of items showed plausibility effects occurring in the ambiguous region, as happened in the full set. We conclude, therefore, that effects in the main analyses were not due to the four items that behaved differently in the fragment completion task. In conclusion, the plausibility effect demonstrates that the processor adopted the NP analysis, even though the verbs were PP-preference. Hence the critical finding of Experiments 1 and 2 generalised to a natural reading environment and could therefore not have been affected by strategies associated with the self-paced reading task. The absence of a reversal of the plausible-implausible contrast is somewhat surprising. Pickering and Traxler (1998) did find this reversal, as did Pickering et al. (2000) in two experiments. However, Traxler and Pickering (1996) did not find a significant reversal, and neither did Pickering et al. (2000) in one experiment.

GENERAL DISCUSSION

In all three experiments, the effects of plausibility at the critical verb demonstrated that participants formed the unbounded dependency by computing the NP analysis as soon as they encountered this verb. Most critically, they formed this dependency even when the verb was PP-

preference, that is, when it more frequently took a noun phrase object than a prepositional phrase. Experiment 3 demonstrated that these effects held in normal reading, and so could not be in any way dependent on particular strategies associated with self-paced reading. The fact that the same pattern of data occurred in all experiments provides good evidence against the conclusions of Stowe et al. (1991).

The experiments also provided some more limited evidence for the cross-over effect, which is indicative of misanalysis (Pickering & Traxler, 1998). In Experiments 1 and 2, sentences with plausible initial analyses were easier than sentences with implausible initial analyses before disambiguation, but harder after disambiguation (though effects after disambiguation were not significant by items in Experiment 2). These results provide more evidence that participants constructed the NP analysis irrespective of the verb's subcategorisation preference, and also that they committed more strongly to this analysis when it was plausible than implausible.

Our results rule out the lexical-frequency account, as this account claims that the initial decision about whether to form the unbounded dependency is dependent on subcategorisation preferences alone. They also rule out an account on which the decision is based on preferences defined by the combination of subject plus verb, as the results remained significant on the subset of items where the preferences did not change. Both of these accounts are deterministic in that the same preference is always adopted for the same ambiguity. But in an alternative 'probabilistic' account, the processor adopts the alternative analyses in relation to their frequency. This simplest account would directly relate frequencies to probabilities, so that if the NP analysis is three times as frequent as the PP analysis, then the processor will adopt the NP analysis three times as often as its alternative. This account is extremely unlikely, because it would predict reduced effects in the PP-preference conditions versus the NP-preference conditions of Experiments 1 and 2. In fact, the effects for both pairs of conditions were almost identical in Experiment 1, and actually tended to be (non-significantly) stronger for the PP-preference conditions in Experiment 2. Obviously, we cannot entirely rule out an account on which the processor adopts the PP analysis on a small fraction of trials in the PP-preference condition, but there is nothing in our data to suggest that this position is correct.[2]

[2] Note that Traxler et al. (1998) and Van Gompel et al. (2001) proposed a ''race-based account'' of attachment ambiguities, which appears compatible with a probabilistic model. However, they specifically argued that ambiguity resolution is probabilistic when two alternatives are roughly equally likely, and where both involve attachment (e.g., high vs. low attachment of relative clauses into complex noun phrases). It is apparent from our data that unbounded dependencies of the kind considered here are resolved in a very different way.

Overall, the results demonstrate that the lexical-frequency account of the processing of unbounded dependencies is not correct. Hence we can be certain that the processor does not decide whether to form unbounded dependencies on the basis of subcategorisation preferences alone.

Constraint-based accounts

Constraint-based accounts assume that subcategorisation preferences constitute an important constraint on the decision about which analysis to favour during initial processing (MacDonald et al., 1994). Moreover, theorists within this tradition have used experiments that test the role of subcategorisation preferences in initial processing as a source of evidence in support of such accounts (Garnsey et al., 1997; Trueswell et al., 1993). It follows that if subcategorisation preferences do not play a central role in such processing decisions, that the empirical basis of constraint-based accounts is thereby weakened.

Constraint-based accounts are, however, more general than lexical-frequency accounts, in that the factors that determine initial preferences are not limited to the frequency with which a verb is used on a given analysis. Instead, a broader range of contextual information can be relevant, including the whole of the prior context. In our whole fragment norming, completions did not accord with the subcategorisation preferences that came from verb norming, but rather depended on the plausibility of the NP analysis. Assuming that the predictions of constraint-based accounts can be predicted largely or entirely from completion preferences, then their predictions differ from those of the lexical frequency account. Specifically, participants should favour the PP analysis when the NP analysis is implausible, and the NP analysis when the NP analysis is plausible.[3]

Are our findings compatible with these predictions? If they were correct, the reason that the sentences with implausible NP analyses were initially harder than the sentences with plausible NP analyses would not be because participants initially adopted the NP analysis, irrespective of lexical preferences. Instead, it would have to be because the PP analysis was more complex than the NP analysis, and that this difference emerged right after encountering the critical verb. We know of no reason to believe that this is correct. For example, Rayner and Duffy (1986) found no effects of verb complexity on reading-time in an eye-tracking study (even though there

[3] Some constraint-based accounts may not make these predictions, because lexical preferences play a role that is not entirely subsumed by preferences due to the entire prior context (McRae et al., 1998; Spivey & Tanenhaus, 1998). Their predictions for our sentences cannot be assessed without extensive modelling.

were effects of lexical ambiguity on nouns). However, it is impossible to discount this possibility entirely.

Constraint-based accounts also face another problem with respect to our data. A central tenet of such accounts is that ambiguities are resolved by competition, and that competition is greater when different analyses are activated to a similar extent than when one analysis is much favoured. For example, Spivey-Knowlton and Sedivy (1995) stated that ''processing delays are a manifestation of direct competition between opposing alternatives ... Near equal activation levels of the two alternatives will result in lengthy competition, hence greatly slowed reading at that point of ambiguity'' (p. 260). During the ambiguous region in Experiments 1–3, the conditions with the plausible NP analysis are read faster than the conditions with the implausible NP analysis. According to constraint-based accounts, this must be because there is more competition in the implausible than the plausible conditions. This is correct for the conditions using NP-preference verbs, because both frequency and plausibility support the NP analysis in the plausible condition (so there is little competition), but frequency and plausibility are in conflict for the implausible condition (so there is more competition). But it is incorrect for the conditions using PP-preference verbs, because frequency and plausibility are in conflict in the plausible condition and are not in conflict in the implausible condition. In other words, the principle of competition, which is central to constraint-based accounts, predicts precisely the opposite pattern of results for PP-preference verbs to that actually found. Pickering et al. (2000) reached similar conclusions for syntactic ambiguities that did not involve unbounded dependencies in their study. Moreover, the pattern of results reported by Boland et al. (1995) support these conclusions. Our conclusions are also compatible with findings that globally ambiguous sentences with two analyses of approximately equal activation are processed more easily than disambiguated sentences (Traxler, Pickering, & Clifton, 1998; Van Gompel, Pickering, & Traxler, 2001), contrary to the predictions of constraint-based accounts.

First-resort accounts

In contrast, the results are consistent with the first-resort strategy, because that account predicts that the processor adopts the NP analysis initially, irrespective of frequency. Thus they are compatible either with the Active filler strategy (e.g., Frazier & Clifton, 1989) or a gap-free alternative (Pickering, 1994). More generally, they lend support to accounts of syntactic ambiguity resolution that are not primarily frequency driven, in accordance with other recent empirical findings (Kennison, 2001; Pickering et al., 2000).

One reason to question whether the first-resort strategy entirely explains the processing of locally ambiguous unbounded dependencies concerns cases where there is no subcategorisation ambiguity. In a sentence like *That's the diver that the coach persuaded the pupils to watch before the tournament*, the verb *persuaded* takes a noun-phrase object and a clause. On one analysis, the filler *the diver* can be treated as the object of *persuaded*; but on the ultimately correct analysis, it serves as the object of the verb *watch* within the embedded clause. Two recent studies found no evidence of plausibility effects on *persuaded* (by comparing *diver* with *event*; it is possible to persuade a diver but not an event), even though they found reversed plausibility effects downstream (i.e. with the implausible condition being easier to process) (Boland et al., 1995; Pickering & Traxler, 2001). The results do not appear compatible with the first-resort strategy, as that account predicts that *the diver* should initially be attached as the object of *persuaded*. However, Pickering and Traxler (2001) suggested that the lack of an effect might possibly be because most of the verbs in those experiments required an animate object, and the processor might only form an unbounded dependency if a minimal semantic feature check was passed. The 'implausible' conditions in Pickering and Traxler (2001) normally used inanimate objects. In the current experiments, most of the plausible and implausible fillers for the PP-preference items did not differ in animacy (in 11 out of 14 items), so, on this account, the processor would initially form the unbounded dependency in both conditions, in accord with the first resort strategy.

In conclusion, the experiments in this paper indicate that the processor forms and interprets unbounded dependencies in an extremely incremental manner. More specifically, it forms the dependency immediately even if subcategorisation information indicates that the dependency is unlikely to be correct. Thus, we can rule out the lexical-frequency account, and suggest that lexical frequency information is unlikely to play a significant role in the resolution of such ambiguities.

Manuscript received May 2002
Revised manuscript received December 2002

REFERENCES

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

Adams, B.C., Clifton, C. Jr., & Mitchell, D.C. (1998). Lexical guidance in sentence processing? *Psychonomic Bulletin & Review*, *5*, 265–270.

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

Boland, J.E., Tanenhaus, M.K., Garnsey, S.M., & Carlson, G.N. (1995). Verb argument structure in parsing and interpretation: Evidence from wh-questions. *Journal of Memory and Language*, *34*, 413–432.

Bourdages, J. (1992). Parsing complex NPs in French. In H. Goodluck & M. Rochemont (Eds.), *Island constraints: Theory, acquisition and processing* (pp. 61–87). Dordrecht: Kluwer.

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

Chomsky, N. (1981). *Letters on government and binding*. Dordrecht: Foris.

Clifton, C. Jr., & Frazier, L. (1989). Comprehending sentences with long distance dependencies. In G. Carlson & M. Tanenhaus (Eds.), *Linguistic structure in language processing* (pp. 273–317). Dordrecht: Kluwer.

Cohen, J.D., MacWhinney, B., Flatt, M., & Provost, J. (1993). Psyscope: A new graphic interactive environment for designing psychological experiments. *Behavioral Research Methods, Instruments, & Computers*, *25*, 257–271.

Crain, S., & Fodor, J.D. (1985). How can grammars help parsers? In D.R. Dowty, L. Karttunen, & A. Zwicky (Eds.), *Natural language parsing* (pp. 94–128). Cambridge: Cambridge University Press.

Crocker, M.W. (1996). *Computational psycholinguistics: An interdiciplinary approach to the study of language*. Dordrecht: Kluwer.

De Vincenzi, M. (1991). *Syntactic parsing strategies in Italian*. Dordrecht: Kluwer.

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

Ferreira, F., & Henderson, J. (1990). Use of verb information in syntactic parsing: Evidence from eye movements and word-by-word self-paced reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 555–568.

Fodor, J.D. (1978). Parsing strategies and constraints on transformations. *Linguistic Inquiry*, *9*, 427–473.

Fodor, J.D. (1989). Empty categories in sentence processing. *Language and Cognitive Processes*, *4*, 155–209.

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

Frazier, L. (1979). *On comprehending sentences: Syntactic parsing strategies*. Ph.D. Thesis, University of Connecticut.

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII* (pp. 559–586). Hove, UK: Lawrence Erlbaum Associates Ltd.

Frazier, L., & Clifton, C. Jr. (1989). Successive cyclicity in the grammar and parser. *Language and Cognitive Processes*, *4*, 93–126.

Frazier, L., & Flores D'Arcais, G.B. (1989). Filler driven parsing: A study of gap filling in Dutch. *Journal of Memory and Language*, *28*, 331–344.

Freedman, S., & Forster, K.I. (1985). The psychological status of overgenerated sentences. *Cognition*, *19*, 101–131.

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

Garnsey, S.M., Tanenhaus, M.K., & Chapman, R. (1989). Evoked potentials and the study of sentence comprehension. *Journal of Psycholinguistic Research*, *18*, 51–60.

Garrod, S.C., Freudenthal, D., & Boyle, E. (1994). The role of different types of anaphor in the on-line resolution of sentences in a discourse. *Journal of Memory and Language*, *33*, 39–68.

Gibson, E., & Hickok, G. (1993). Sentence processing with empty categories. *Language and Cognitive Processes*, *8*, 147–161.

Gorrell, P. (1995). *Syntax and parsing.* Cambridge: Cambridge University Press.

Hare, M., McRae, K., & Elman, J.L. (2003). Sense and structure: Meaning as a determinant of verb subcategorisation preferences. *Journal of Memory and Language*, *48*, 281–303.

Kennison, S.M. (2001). Limitations on the use of verb information during sentence comprehension. *Psychonomic Bulletin & Review*, *8*, 132–138.

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

MacDonald, M.C. (1994). Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes*, *9*, 157–201.

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

McElree, B., & Griffith, T. (1998). Structural and lexical constraints on filling gaps during sentence comprehension: A time-course analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 432–460.

McKoon, G., Albritton, D., & Ratcliff, R. (1996). Sentential context effects on lexical decisions with a cross-modal instead of an all-visual procedure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1494–1497.

McKoon, G., & Ratcliff, R. (1994). Sentential context and on-line lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 1239–1243.

McKoon, G., Ratcliff, R., & Ward, G. (1994). Testing theories of language processing: An empirical investigation of the on-line lexical decision task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 1219–1228.

McRae, K., Spivey-Knowlton, M.J., & Tanenhaus, M.K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language*, *38*, 283–312.

Mecklinger, A., Schriefers, H., Steinhauer, K., & Friederici, A.D. (1995). Processing relative clauses varying on syntactic and semantic dimensions: An analysis with event-related potentials. *Memory & Cognition*, *23*, 477–494.

Mitchell, D.C. (1987). Lexical guidance in human parsing: Locus and processing characteristics. In M. Coltheart (ed.), *Attention and performance XII* (pp. 601–618). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Mitchell, D.C., Cuetos, F., Corley, M.M.B., & Brysbaert, M. (1995). Exposure-based models of human parsing: Evidence for the use of coarse-grained (non-lexical) statistical records. *Journal of Psycholinguistic Research*, *24*, 469–488.

Nicol, J., Fodor, J.D., & Swinney, D.A. (1994). Using cross-modal lexical decision tasks to investigate sentence processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 1229–1238.

Nicol, J., & Pickering, M.J. (1993). Processing syntactically ambiguous sentences: Evidence from semantic priming. *Journal of Psycholinguistic Research*, *22*, 207–237.

Nicol, J., & Swinney, D. (1989). The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research*, *18*, 5–19.

Pickering, M.J. (1993). Direct association and sentence processing: A reply to Gorrell and to Gibson and Hickok. *Language and Cognitive Processes*, *8*, 163–196.

Pickering, M.J. (1994). Processing local and unbounded dependencies: A unified account. *Journal of Psycholinguistic Research*, *23*, 323–352.

Pickering, M.J., & Barry, G. (1991). Sentence processing without empty categories. *Languge and Cognitive Processes*, *6*, 229–259.

Pickering, M.J., Barton, S., & Shillcock, R.C. (1994). Unbounded dependencies, island constraints and processing complexity. In C. Clifton, Jr., L. Frazier, & K. Rayner (Eds.),

Perspectives on sentence processing (pp. 199–224). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Pickering, M.J., & Traxler, M.J. (1998). Plausibility and recovery from garden paths: An eye-tracking study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 940–961.

Pickering, M.J., & Traxler, M.J. (2001). Strategies for processing unbounded dependencies: Lexical information and verb-argument assignment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 1401–1410.

Pickering, M.J., Traxler, M.J., & Crocker, M.W. (2000). Ambiguity resolution in sentence processing: Evidence against likelihood. *Journal of Memory and Language*, *43*, 447–475.

Pollard, C., & Sag, I.A. (1993). *Head-driven phrase structure grammar*. Stanford, CA and Chicago, IL: CSLI and University of Chicago Press.

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.

Rayner, K., & Duffy, S.A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition*, *14*, 191–201.

Ross, J.R. (1967). *Constraints on variables in syntax*. Unpublished Ph.D. dissertation, MIT.

Schriefers, H., Friederici, A.D., & Kühn, K. (1995). The processing of locally ambiguous relative clauses in German. *Journal of Memory and Language*, *34*, 499–520.

Spivey, M., & Tanenhaus, M.K. (1998). Syntactic ambiguity resolution in discourse: Modeling the effects of referential context and lexical frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1521–1543.

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

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

Stowe, L.A., Tanenhaus, M.K., & Carlson, G.M. (1991). Filling gaps on-line: Use of lexical and semantic information in sentence processing. *Languge and Speech*, *34*, 319–340.

Tanenhaus, M.K., Carlson, G., & Trueswell, J.C. (1989). The role of thematic structures in interpretation and parsing. *Language and Cognitive Processes*, *4*, 211–234.

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

Traxler, M.J. (2002). Plausibility and subcategorisation preference in children's processing of temporarily ambiguous sentences: Evidence from self-paced reading. *Quarterly Journal of Experimental Psychology*, *55A*, 75–96.

Traxler, M.J., Bybee, M., & Pickering, M.J. (1997). Influence of connectives on language comprehension: Eye-tracking evidence for incremental interpretation. *Quarterly Journal of Experimental Psychology*, *50A*, 481–497.

Traxler, M.J., & Pickering, M.J. (1996). Plausibility and the processing of unbounded dependencies: An eye-tracking study. *Journal of Memory and Language*, *35*, 454–475.

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

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

Trueswell, J., 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.

Tyler, L.K., & Marslen-Wilson, W.D. (1977). The on-line effects of semantic context on syntactic processing. *Journal of Verbal Learning and Verbal Behavior*, *16*, 683–692.

Van Gompel, R.P.G., & Pickering, M.J. (2001). Lexical guidance in sentence comprehension: A reply to Adams, Clifton, and Mitchell (1998). *Psychonomic Bulletin & Review*, *8*, 851–857.

Van Gompel, R.P.G., Pickering, M.J., & Traxler, M.J. (2001). Making and revising syntactic commitments: Evidence against current constraint-based and two-stage models. *Journal of Memory and Language*, *45*, 283–307.

APPENDIX

Experimental items

All items were used in Experiments 1 and 2. Only the PP-preference items were used in Experiment 3. The words within parentheses were included in Experiments 1 and 3 and excluded in Experiment 2. The noun before the '/' was employed in the plausible sentences; the noun after the '/' was employed in the implausible sentences.

PP-preference items

That's the cat/car that the dog worried [compulsively] about after going to the vet because of an injury.
Those are the lines/props that the actor spoke [briefly] about to the acting coach after the rehearsal.
That's the plane/truck that the pilot landed [carefully] behind in the fog at the airport.
That's the secret/record that Sally blabbed [accidentally] about to John this morning.
That's the message/traitor that the courier communicated [secretly] about to the spy in the black hat.
That's the lake/page that the guy fished [repeatedly] for in his atlas during fishing season.
That's the stick/river that the child pointed [excitedly] toward this morning down by the shore.
That's the sermon/sinner that the reverend preached [fervently] about in church on Sunday.
That's the car/pit that Dale Earnhardt raced [skillfully] past near the end of the race last week.
That's the suspect/weapon that the policeman searched [rapidly] for in the alley after the murder.
Those are the insults/animals that the protestors shouted [loudly] about at the meeting in the dormitory.
That's the highway/factory that the trucker traveled [rapidly] toward after picking up a load from the mine.
That's the case/bomb that the lawyer argued [heatedly] about to the judge during the trial.
That's the oath/tank that the soldier swore [emphatically] about to the sergeant in the mess hall.

NP-preference items

That's the general/country that the soldier killed [enthusiastically] for during the war in Korea.
That's the teacher/college that the student asked [repeatedly] about during orientation last summer.

That's the worker/gadget that the manager called [occasionally] about after the accident at the factory.
That's the noise/field that the fisherman heard [repeatedly] about at the bar after fishing trips.
That's the blanket/village that the girl hid [secretly] inside after being chased by a dog.
That's the class/house that the architect instructed [at length] about to the graduate students in the seminar.
That's the businessman/corporation that the ex-con kidnapped [reluctantly] for after being released from prison.
That's the plumber/tractor that the contractor paid [excessively] for this summer because of the union contract.
That's the machine/patient that the surgeon operated [carefully] on at the hospital on Monday.
That's the article/burglar that the editor read [meticulously] about in the office before the meeting.
That's the lesson/carpet that the teaching assistant taught [inexpertly] about to the undergraduates in the art school.
That's the rock/dock that the boy threw [thoughtlessly] at before being scolded by his mother.
That's the story/fight that the old man told [frequently] about to the residents of the nursing home.
That's the poem/city that the young woman wrote [emotionally] about to herself in her diary.