Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Memory and Language 54 (2006) 389–406
Journal of Memory and Language
www.elsevier.com/locate/jml

Heavy NP shift is the parser's last resort: Evidence from eye movements ☆

Adrian Staub [a,*], Charles Clifton Jr. [a], Lyn Frazier [b]

[a] *Department of Psychology, University of Massachusetts, Amherst, MA 01003, USA*
[b] *Department of Linguistics, University of Massachusetts, Amherst, MA 01003, USA*

Received 19 July 2005; revision received 2 December 2005
Available online 2 February 2006

Abstract

Two eye movement experiments explored the roles of verbal subcategorization possibilities and transitivity biases in the processing of heavy NP shift sentences in which the verb's direct object appears to the right of a post-verbal phrase. In Experiment 1, participants read sentences in which a prepositional phrase immediately followed the verb, which was either obligatorily transitive or had a high transitivity bias (e.g., *Jack praised/watched from the stands his daughter's attempt to shoot a basket*). Experiment 2 compared unshifted sentences to sentences in which an adverb intervened between the verb and its object, and obligatorily transitive verbs to optionally transitive verbs with widely varying transitivity biases. In both experiments, evidence of processing difficulty appeared on the material that intervened between the verb and its object when the verb was obligatorily transitive, and on the shifted direct object when the verb was optionally transitive, regardless of transitivity bias. We conclude that the parser adopts the heavy NP shift analysis only when it is forced to by the grammar, which we interpret in terms of a preference for immediate incremental interpretation.

Keywords: Sentence comprehension; Parsing; Heavy NP shift; Argument structure; Eye movements

Consider the comprehension of sentences in which a direct object is separated from the verb by intervening material, as in (1) below:

(1) Lucy ate t_i with a fork [the extremely delicious, bright green broccoli]$_i$.

This construction is known as heavy NP shift. According to the classic syntactic account (Ross, 1967; but cf. Kayne, 1998; Rochemont & Culicover, 1997), heavy NP shift involves a movement operation that displaces the verb's direct object from its underlying position adjacent to the verb, where it receives both accusative case and a thematic role, to an adjoined position to the right of other verbal arguments and/or adjuncts. As in other movement operations, a trace of the moved constituent remains in the pre-movement position. Notably, heavy NP shift is typically able to

☆ This research was supported by National Institutes of Health Grant HD-18708 to the University of Massachusetts and by a University of Massachusetts Graduate School Fellowship to the first author. Portions of this research were presented at the Eighteenth Annual CUNY Conference on Human Sentence Processing, Tuscon, Arizona, April 2005. We thank Keith Rayner for his helpful comments.
* Corresponding author. Fax: +1 413 545 0996.
E-mail address: astaub@psych.umass.edu (A. Staub).

0749-596X/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.jml.2005.12.002

apply only when the direct object noun phrase is relatively long or "heavy" relative to the intervening constituent, as evidenced by the comparative awkwardness of (2a) and the unacceptability of (2b):

(2a) Lucy ate t_i with a fork [the broccoli]$_i$.
(2b) Lucy ate t_i with a fork [peas]$_i$.

To our knowledge, there have been no published experimental studies of the comprehension of heavy NP shift constructions. (We know of only one experimental study of heavy NP shift production: Stallings, MacDonald, & O'Seaghdha, 1998.) This is surprising for two reasons. First, heavy NP shift is relatively common. Wasow (1997) found that in a corpus of printed materials, heavy NP shift accounted for between 5 and 10% of all sentences with direct objects, with this rate varying by verb type (in a manner to be discussed below). In a separate corpus study focusing on certain idiomatic constructions, Wasow found heavy NP shift rates of around 50%.

Second, and perhaps more importantly, heavy NP shift has much in common with *wh*-movement, the processing of which has been studied in great detail (e.g., Aoshima, Phillips, & Weinberg, 2004; Boland, Tanenhaus, Garnsey, & Carlson, 1995; Clifton & Frazier, 1989; Crain & Fodor, 1985; Fodor, 1978, 1989; Frazier & Clifton, 1989; Garnsey, Tanenhaus, & Chapman, 1989; Pickering & Traxler, 2003; Stowe, 1986; Traxler & Pickering, 1996). *Wh*-movement occurs in both questions and relative clauses, as in (3a) and (3b), respectively:

(3a) [Which boy]$_i$ did the teacher punish t_i for his bad behavior?
(3b) The teacher punished the boy [whom]$_i$ she had asked t_i to stop running around.

Like heavy NP shift, *wh*-movement can displace a direct object noun phrase (or other argument of a verb) from its underlying position, leaving in this position a trace of movement (in psycholinguistic terminology, a *gap*). In comprehending a sentence in which either heavy NP shift or *wh*-movement has taken place, the reader or listener must associate the moved constituent (or *filler*) with the underlying position in which this constituent is assigned a thematic role by a verb. An important difference between these constructions, however, is that *wh*-movement results in a word order in which the direct object or other argument has moved closer to the beginning of the sentence, while heavy NP shift displaces the direct object to a position later in the sentence. As a result, the order in which a reader or listener encounters the filler and the corresponding gap is reversed in these two constructions.

In sentences that involve *wh*-movement, the filler is usually easily identifiable, but the gap is not signaled unambiguously. It is possible for the parser to posit a gap in a position that turns out to be incorrect, or to overlook the true gap site. Consider the following modification of (3a) above:

(4) [Which boy]$_i$ did the teacher punish the girls for teasing t_i?

As (3a) indicates, the gap corresponding to *which boy* could occur after *punish*; but as (4) indicates, this gap could instead appear later in the sentence, after *teasing*. A reader or listener who does posit a gap site after *punish* will have to reanalyze upon encountering *the girls* in (4). On the other hand, a reader or listener who does not initially posit a gap site after *punish* will have to do so after encountering *for his* in (3a), since *punish* requires a direct object.

In an influential paper, Fodor (1978) distinguished three strategies or principles that the processor could use to identify the gap location. The processor could adopt a highly conservative *last resort* strategy, positing a gap only when the grammar requires one. On the other hand, the processor could adopt the liberal *first resort* strategy of positing a gap in the first grammatically licensed location. Finally, the parser could posit a gap following a particular verb only after assessing whether the verb frequently takes an argument of the same type as the filler (e.g., a noun phrase or prepositional phrase), and noting whether an argument of this type appears following the verb. If the verb does tend to take an argument of the relevant type, and one does not immediately appear, then the processor posits a gap after the verb. This is the *lexical expectation* strategy.

The last resort strategy predicts that the parser will never posit a gap when an alternative analysis is available. Fodor notes that this prediction is inconsistent with clear intuitions about processing difficulty, and research using on-line measures such as eyetracking and self-paced reading has consistently provided evidence that readers do not follow the last resort strategy (Crain & Fodor, 1985; Stowe, 1986). In fact, a fair amount of evidence now supports the first resort strategy over the lexical expectation strategy (e.g., Frazier & Clifton, 1989; Pickering & Traxler, 2003; Traxler & Pickering, 1996; but cf. Boland et al., 1995). For example, Pickering and Traxler (2003) tracked participants' eye movements as they read sentences like (5a and 5b):

(5a) That's the plane that the pilot landed behind in the fog at the airport.
(5b) That's the truck that the pilot landed behind in the fog at the airport.

Participants took longer to read the region that included the verb *landed* in (5b), and took longer to read the next region in (5a). This suggests that readers were interpreting *the plane* or *the truck* as the direct object of *landed*, even though the verb is less frequently used with a noun phrase complement than with a prepositional phrase complement. In (5a), this analysis can no longer be maintained when the reader reaches *in the fog*; in (5b), the implausibility of landing a truck causes the reader to reanalyze almost immediately.

In this article, we present two experiments designed to examine the question of when, in the course of processing a heavy NP shift sentence, the processor first constructs the heavy NP shift analysis. It is possible to distinguish first resort, last resort, and lexical expectation strategies that are analogues to Fodor's (1978) three strategies for the processing of *wh*-movement. A processor that adopts the first resort strategy will rank the option of heavy NP shift above any other option; i.e., in (1) above it will posit a gap after the verb *ate* even before checking whether a direct object follows immediately in the input. Intuitively this strategy seems quite unlikely, since it suggests that integrating a direct object after *ate* (e.g., *Lucy ate the broccoli with a fork*) should actually induce difficulty because the direct object position is already filled by a trace of movement. More plausible is a lexical expectation strategy, which predicts that the parser will posit a trace of a shifted noun phrase if the verb is relatively transitive-biased and a direct object does not immediately appear. If this is the parser's strategy, then the heavy NP shift analysis will not be constructed for (1) until the reader or listener begins to process *with a fork*. Even then, the parser will construct the heavy NP shift analysis only if the verb *eat* is sufficiently transitive-biased. If *eat* is not sufficiently transitive-biased, the parser will assume, upon encountering *with a fork*, that *eat* is occurring in an intransitive frame. It will then be forced to construct the heavy NP shift analysis upon encountering the shifted noun phrase itself (*the extremely delicious, bright green broccoli*).

Finally, the last resort strategy predicts that the parser will posit a trace of a shifted noun phrase only when the grammar forces the heavy NP shift analysis. Crucially for our experiments, the point at which the grammar forces this analysis varies based on the subcategorization possibilities associated with the verb. If the verb is obligatorily transitive, the heavy NP shift analysis is forced when the reader or listener encounters the material intervening between the verb and the direct object. For example, if the verb *devour* were substituted for *eat* in (1), the heavy NP shift analysis would be required upon reaching the prepositional phrase *with a fork*, since there is no grammatical continuation of *Lucy devoured with a fork* other than a heavy NP shift construction. But if the verb is optionally transitive, like *eat*, non-heavy NP shift continuations are possible, e.g., *Lucy ate with a fork rather than a spoon*. In this case, the heavy NP shift analysis is not forced until the shifted noun phrase itself is encountered.

The differences between the predictions of the lexical expectation and last resort strategies center on the role of verb subcategorization information. According to the lexical expectation strategy, the point at which the parser constructs the heavy NP shift analysis should vary depending on verb subcategorization *bias*, with high-transitive verbs increasing the likelihood of an early (i.e., on the immediately post-verbal material) adoption of this analysis. But according to the last resort strategy, only the possibility of an intransitive frame matters, not the frequency of this frame. If the verb occurs only in a transitive frame, the parser will construct the heavy NP shift analysis on the immediately post-verbal material; but if the verb allows an intransitive frame, the parser will not construct the heavy NP shift analysis until it reaches the shifted noun phrase itself. In other words, the two strategies make different predictions for verbs that are transitive-biased but not obligatorily transitive: the lexical expectation strategy predicts that with these verbs, the parser will construct the heavy NP shift analysis on the material that intervenes between the verb and direct object, while the last resort strategy predicts that the parser will not construct the heavy NP shift analysis until later.

We assume that the first resort strategy for processing *wh*-movement arises from a general preference to construct syntactic analyses in a manner that maximizes immediate incremental interpretation (Clifton & Frazier, 1989; Crocker, 1996; Frazier, 1987; Pickering, 1994). We think that the same general preference argues for the last resort strategy in the processing of heavy NP shift constructions. A preference for maximizing incremental interpretation is evident in the processing of a range of structural ambiguities, resulting in, for example, the well-attested preferences for Minimal Attachment and Late Closure (Frazier, 1978; Frazier & Rayner, 1982). In a *wh*-movement construction, the filler cannot be assigned a role in the sentence's meaning until it is related to a particular verb (or other phrasal head) as a theme, goal, instrument, etc. Incremental interpretation is maximized by locating the underlying position of the moved constituent as quickly as possible. In a heavy NP shift construction, on the other hand, the interpretation of the verb phrase is speeded if the parser avoids positing a trace of a moved direct object. In (1) above, adopting an intransitive frame upon encountering *with a fork* allows the reader or listener to interpret this phrase as modifying *eat*. By contrast, adopting the heavy NP shift analysis requires the comprehender to delay semantic interpretation, because on this analysis *with a fork* is modifying the eating of something specific, but it is not yet clear what.

Before presenting the experiments in detail, we note the relevance of the issue we are investigating to a proposal by Wasow (1997). Wasow performed several corpus studies to explore the circumstances under which speakers and writers produce sentences with a shifted word order. Specifically, he investigated whether speakers and writers produce the heavy NP shift word order for their own benefit, or whether they take into consideration the needs of the listener or reader. He assumed that the heavy NP shift word order is especially difficult for the comprehender when the verb allows an intransitive frame, since in these circumstances the parser could be temporarily misled into adopting this frame. If language producers take this fact into consideration, heavy NP shift constructions should appear more frequently with obligatorily transitive verbs than with optionally transitive verbs. He found, however, that in both printed materials and spoken language heavy NP shift is in fact more common with optionally transitive verbs than with obligatorily transitive verbs. Wasow suggested that by using the heavy NP shift word order with optionally transitive verbs, producers are leaving their options open, since they can decide relatively late in the sentence whether or not to include a direct object.

Our hypothesis that heavy NP shift is the parser's last resort is quite consistent with Wasow's assumption that listeners can be misled into adopting an intransitive frame. In fact, the last resort strategy predicts that listeners will always adopt an intransitive frame, when one is available. We also think that, as Wasow implies, revising this initial analysis in favor of the heavy NP shift analysis is likely to induce processing difficulty.

Unlike Wasow, however, we suspected that the heavy NP shift word order may also induce processing difficulty when the verb is obligatorily transitive. When material other than a direct object arrives after an obligatorily transitive verb (e.g., *Lucy devoured with a fork*), the parser has no choice but to construct the heavy NP shift analysis. However, the fact that the verb is obligatorily transitive may cause readers or listeners initially to attempt to attach the post-verbal material as the direct object, and to have to reanalyze. This is especially likely to be the case if the parser has a predictive component, enabling it to build the structure corresponding to the direct object upon processing the verb (e.g., Frazier & Fodor, 1978; Kimball, 1973). We hypothesized, in short, that processing heavy NP shift sentences is likely to be difficult whether the verb is obligatorily or optionally transitive; we designed our experiments to test the idea that this difficulty would appear on the material that intervenes between verb and direct object when the verb is obligatorily transitive, and on the shifted noun phrase when the verb is optionally transitive.

Experiment 1

Experiment 1 was designed to test the predictions of the last resort strategy for the processing of heavy NP shift against the predictions of the first resort and lexical expectation strategies. We had participants read heavy NP shift sentences in which we varied the subcategorization possibilities associated with the verb, and we recorded their eye movements while they read. Eye movements are arguably the best indicator of moment-to-moment processing during reading (Rayner, 1998). We presented sentences like the following:

(6a) Jack praised from the stands his daughter's attempt to shoot a basket.
(6b) Jack watched from the stands his daughter's attempt to shoot a basket.

In (6a), the main verb (*praised*) requires a direct object. Consequently, the reader must adopt the heavy NP shift analysis to attach the immediately post-verbal prepositional phrase (*from the stands*). In (6b), on the other hand, the verb (*watched*) is strongly transitive-biased, but allows an intransitive frame; the sentence could continue with, for example, an adjunct clause after the post-verbal prepositional phrase, e.g., *Jack watched from the stands as his daughter tried to shoot a basket*. In (6b), it is the shifted direct object noun phrase itself (*his daughter's attempt to shoot a basket*) that alerts the reader that the sentence requires a heavy NP shift analysis.

The last resort strategy predicts that readers will, in fact, adopt the heavy NP shift analysis at these two different points in the two sentences. Consistent with the assumption discussed above that constructing this analysis is likely to induce processing difficulty, we made two specific predictions regarding the pattern of data. First, we expected to see longer reading times and/or more regressive eye movements on the prepositional phrase in (6a) than in (6b). Second, we expected to see shorter reading times and/or fewer regressions on the shifted noun phrase in (6a) than in (6b). In short, we predicted a processing tradeoff, with difficulty arising in each sentence at the point at which the processor has no option but to construct the heavy NP shift analysis. This point is relatively early when the verb is obligatorily transitive, and relatively late when the verb is optionally transitive.

The first resort and lexical expectation strategies do not predict differences between (6a) and (6b). The first resort strategy predicts that the heavy NP shift analysis is the parser's initial analysis, regardless of verb subcategorization possibilities. Therefore, it predicts that this analysis should induce difficulty in neither (6a) nor (6b). As noted above, we regarded the first resort strategy as implausible, and considered the lexical expectation strategy to be the serious competitor to the last

resort strategy. The lexical expectation strategy predicts that the parser will construct the heavy NP shift analysis prior to reaching the shifted direct object in both (6a) and (6b). In (6a), this analysis is required by the grammar. In (6b), the verb (*watched*) is transitive-biased, and according to the lexical expectation strategy, this will induce the parser to adopt the heavy NP shift analysis on the prepositional phrase. To distinguish the lexical expectation and last resort strategies, we made sure that all of our optionally transitive verbs were in fact transitive-biased.

Method

Participants

Thirty-six native speakers of American English, who were students at the University of Massachusetts, were given course credit or were paid $5 to participate in the experiment. All had normal or corrected-to-normal vision, and all were naïve to the purpose of the experiment.

Materials

We constructed 24 sets of sentences like (6a) and (6b). In each sentence, a prepositional phrase appeared immediately after the verb, followed by a direct object noun phrase. In all but one item, the prepositional phrase was a locative or temporal adjunct of the verb. In the remaining item, the prepositional phrase was an argument of the verb in both sentences. The noun phrase was made relatively long (or "heavy") by including a relative clause, a prepositional phrase modifier, or an argument of the head noun. The only difference between the two versions was the choice of verb. In the (a) version, the verb was obligatorily transitive; we (and several independent raters) judged that a direct object had to appear after the prepositional phrase in order for the sentence to be grammatical. We refer to this as the *obligatory* condition. In the (b) version, the verb allowed for a range of possible continuations. We refer to the this as the *optional* condition. The full set of materials is provided in Appendix A.

The 24 sets of sentences used 12 different verb pairs, with each verb pair appearing in two sets of sentences. Each verb pair was matched for length, and the obligatorily and optionally transitive verbs did not differ significantly in frequency, based on the Francis and Kučera (1982) norms. The obligatorily transitive verbs had a mean $\log_n$ frequency of 3.93, $SD = 1.27$, and the optionally transitive verbs had a mean $\log_n$ frequency of 4.81, $SD = 1.15$. Two lists were constructed so that each participant saw one sentence from each pair, and 12 sentences overall in each of the two conditions. The lists of stimuli were arranged so that each participant saw each verb only once; i.e., if a participant saw the version of (6) that included the obligatorily transitive verb *praised*, he or she would see the version of the other *praised/watched* stimulus set that included the verb *watched*.

The 12 optionally transitive verbs that we used were all transitive-biased. These 12 verbs appeared in both the Connine, Ferreira, Jones, Clifton, and Frazier (1984) production norms and the Gahl, Jurafsky, and Roland (2004) corpus-based norms. We calculated the transitivity bias of each verb as the mean of the scores obtained from these two sources, using scoring categories that were common to both sets of norms. Note, however, that while all the verbs had high transitivity ratings according to both sets of norms, the correlation between the scores obtained from these two sources was very close to zero ($r = .01$), presumably because of the restricted range of transitivity ratings for the verbs we used. Table 1 provides the transitivity biases obtained from each set of norms, as the percentage of the total occurrences of each verb in which the verb was used in a transitive frame, as well as the means of these scores. The mean transitivity biases of the optionally transitive verbs ranged from .66 (*worship*) to .93 (*visit*), with overall mean of .80 and $SD = .07$.

Of the 12 obligatorily transitive verbs that we used, three appeared in neither the Connine et al. nor the Gahl et al. norms, and two others appeared only in the Connine et al. norms. To obtain transitivity bias ratings from a common source for all 12 verbs, we conducted a production norming study with 16 participants. For each verb, participants were asked to write a sentence containing the verb, using the "first sentence that comes to mind." The verbs of interest were intermixed with several optionally transitive and intransitive verbs. We excluded five responses in which the participant used the critical word as a noun, rather than a verb, leaving a total of 187 responses. Of these, six responses fell into categories that were clearly not relevant to the experiment. Four used the verb *bother* in a "negative polarity" construction (e.g., *Don't bother to clean it up.*); this

Table 1
Experiment 1: Transitivity biases of optionally transitive verbs

Verb	Connine et al. (1984)	Gahl et al. (2004)	Overall
worship	.72	.60	.66
govern	.79	.75	.77
attack	.77	.94	.86
watch	.84	.73	.79
leave	.82	.78	.80
visit	.96	.91	.93
follow	.94	.66	.80
push	.94	.64	.79
paint	.89	.89	.89
hear	.64	.78	.71
teach	.83	.79	.81
clean	.80	.89	.84

construction can be ruled out by the reader of our experimental sentences by the time he or she has reached the verb, since the verb has not been preceded by a licensing negative element. In two other cases, the verb was used in a common idiom (e.g., *Carry on*.). A reader of our experimental sentences can rule out such an idiomatic use by the time the post-verbal prepositional phrase is reached. The verb had a direct object noun phrase in 179 of the remaining 181 responses (99%).

We note that exactly half of the experimental items had a definite description (e.g., *the sailors*, *the math teacher*), rather than a name, in subject position. For these 12 items, it was possible in principle for participants to adopt an initial analysis according to which the verb marked the beginning of a reduced relative clause (though in many cases this analysis would have been very implausible). We considered it quite unlikely that participants would actually adopt this analysis, since the reduced relative analysis is strongly dispreferred, in most contexts (e.g., Binder, Duffy, & Rayner, 2001; Ferreira & Clifton, 1986; cf. Spivey & Tanenhaus, 1998). In any event, statistical analyses showed no significant differences between items that did, and did not, temporarily allow a reduced relative analysis, so we do not distinguish between these two groups of items in reporting our results.

The experimental sentences were intermixed with 118 filler sentences of various types, none of which used the heavy NP shift word order. The sentences were presented in an individually randomized order to each participant.

Procedure

Participants were tested individually. Eye movements were recorded using a Fourward Technologies Dual Purkinje Generation 6 eyetracker, which has an angular resolution of less than 10 min of arc. The eyetracker was interfaced with an IBM compatible computer. All sentences in this experiment were displayed on a single line, with a maximum length of 80 characters. While viewing was binocular, only the right eye was monitored. Stimuli were displayed on a 15-in. NEC MultiSync 4FG monitor. Participants were seated 61 cm from the computer screen; at this distance, 3.8 characters subtended 1° of visual angle.

On arrival at the laboratory, participants were given instructions and had a bite bar prepared for them that served to stabilize the head. A calibration routine was performed, and its accuracy was checked after each sentence. Participants were instructed to read the sentences for understanding, and to read at a normal rate. After reading each sentence, the participants pressed a button to remove the sentence. The first eight trials of the experimental session were practice trials. Comprehension was checked on approximately 30% of all trials during the experiment by presenting the participant with a yes/no question. Average accuracy for the comprehension questions was above 85%, with no participant scoring below 75%. The entire experiment lasted approximately 30 min.

Results and discussion

We analyzed three regions in each sentence. The *prepositional phrase region* consisted of the prepositional phrase that appeared immediately following the verb (*from the stands* in 6a and 6b). As discussed above, we predicted elevated reading times and/or an increase in regressions on the prepositional phrase in the obligatory condition compared to the optional condition. The next region, the *noun phrase region*, consisted of the direct object noun phrase up through the head noun (*his daughter's attempt*). This region was three or four words in length, with the exception of one item in which it was five words. We predicted elevated reading times and/or an increase in regressions on the noun phrase in the optional condition compared to the obligatory condition. Finally, we analyzed a *spillover region* that consisted of the remainder of the sentence (*to shoot a basket*).

Four reading time measures were computed: *first fixation duration*, *first pass time* (which is referred to as *gaze duration* when discussing single-word regions), *go-past time*, and *percent regressions* (Rayner, 1998). The first two measures reflect early stages of processing such as lexical access (Reichle, Rayner, & Pollatsek, 2003), but parsing difficulty has also been shown to affect these measures (e.g., Frazier & Rayner, 1982). First fixation duration is simply the duration of the first fixation in a region, whether it is the only fixation in the region or the first of multiple fixations. Though this measure is typically used with single-word regions, we computed it in this experiment because we anticipated that effects of verb type might appear as early as the first fixation on the prepositional phrase and/or the shifted noun phrase. First fixation duration on the spillover region is unlikely to be informative, but we report it in the interest of completeness. First pass time is the sum of all fixations in a region prior to leaving the region for the first time, either to the left or the right. Go-past time is the elapsed time from first fixating the region until the reader leaves the region to the right, including any time spent to the left of the region after a regressive eye movement and any time spent re-reading material in the region before moving on. (Go-past time is also sometimes called *regression path duration*.) Finally, the percent regressions measure gives the probability that a reader makes a regressive eye movement after fixating the region. This measure includes only regressions made during the reader's first pass through the region; it does not include regressions made after re-fixating the region. Effects of syntactic misanalysis are often apparent in the go-past and regression measures. We did not compute

late measures such as second pass time and total time because we were interested primarily in the processing costs incurred at the first encounter with the prepositional phrase and the subsequent noun phrase.

Prior to all analyses, sentences with track losses were excluded (less than 2% of trials). In addition, fixations less than 80 ms in duration, and within one character of the previous or subsequent fixation, were incorporated into this neighboring fixation. The same procedure was used to incorporate fixations less than 40 ms in duration and within three characters of the previous or subsequent fixation. Remaining fixations of less than 80 ms were deleted, as were fixations of longer than 800 ms. It is thought that readers do not extract useful information from fixations shorter than 80 ms (see Rayner & Pollatsek, 1989), and that fixations longer than about 800 ms are likely to reflect track losses. Less than 2% of all fixations were eliminated.

Table 2 presents the participant means on each measure for each of the regions we analyzed, as well as 95% confidence intervals for the differences between condition means (Loftus & Masson, 1994; Masson & Loftus, 2003). For each test of differences between means, we performed analyses of variance (ANOVAs) with participants (F_1) and items (F_2) as random effects variables, and with counterbalancing group as a between-participants or between-items factor (Pollatsek & Well, 1995). We report a difference as significant if both of these ANOVAs rejected the null hypothesis at the .05 level. These F statistics, as well as MinF', are reported in Table 3. We discuss the regions individually below.

Prepositional phrase region

The first fixation on this region was longer in the obligatory condition than in the optional condition (280 ms vs. 269 ms), but this difference was significant only in the participants analysis. First pass time, however, was significantly longer in the obligatory condition by both participants and items (646 ms vs. 608 ms). Go-past time was also significantly longer in the obligatory condition (767 ms vs. 668 ms). Finally, there were more regressions in the obligatory condition (14.3% vs. 7.5%). In sum, differences on the first pass, go-past, and percent regressions measures all supported our prediction of greater processing difficulty on the prepositional phrase when the verb was obligatorily transitive than when it was merely transitive-biased.

Noun phrase region

There was no hint of a first fixation difference on the noun phrase region. First pass time, however, was significantly longer in the optional condition than in the obligatory condition (735 ms vs. 693 ms). On the go-past measure, the numerical trend went in the opposite direction, but the difference between conditions was not significant ($ps > .15$). Finally, there was a significant

Table 2
Experiment 1: Participant Mean Reading Times, in Milliseconds, and Percent Regressions

Measure	Prepositional phrase (*from the stands*)	Noun phrase (*his daughter's attempt*)	Spillover (*to shoot a basket*)
First fixation duration			
Obligatory	280	269	259
Optional	269	270	258
95% CI of difference	1.0 to 21.6	−13.4 to 13.0	−10.0 to 11.4
First pass time			
Obligatory	646	693	795
Optional	608	735	794
95% CI of difference	13.6 to 61.4*	−75.5 to −9.7*	−38.8 to 40.2
Go-past time			
Obligatory	767	845	921
Optional	668	809	936
95% CI of difference	61.6 to 131.6*	−16.6 to 89.2	−87.3 to 55.7
Percent regressions			
Obligatory	14.3	11.4	9.2
Optional	7.5	6.5	8.2
95% CI of difference	2.6 to 11.0*	1.2 to 8.6*	−3.2 to 5.2

Obligatory = *Jack praised from the stands his daughter's attempt to shoot a basket.* Optional = *Jack watched from the stands his daughter's attempt to shoot a basket.*

Note. Reported here are the 95% confidence intervals for the differences between conditions, calculated using the procedure for within-participant confidence intervals described by Loftus and Masson (1994; Masson and Loftus, 2003). Differences that were significant by both participants and items ($p < .05$) are marked with an asterisk (*).

Table 3
Experiment 1: Analysis of variance results by region for effect of condition

Measure	Prepositional phrase	Noun phrase	Spillover
First fixation duration	$F_1 = 4.86$, $p < .05$ $F_2 = 2.17$, $p = .16$ $\mathrm{Min}F'(1,40) = 1.50$, $p = .23$	Fs < 1.5	Fs < 1.5
First pass time	$F_1 = 10.14$, $p < .01$ $F_2 = 4.21$, $p = .05$ $\mathrm{Min}F'(1,39) = 2.97$, $p = .09$	$F_1 = 6.93$, $p < .02$ $F_2 = 4.99$, $p < .05$ $\mathrm{Min}F'(1,48) = 2.90$, $p = .10$	Fs < 1.5
Go-past time	$F_1 = 32.80$, $p < .01$ $F_2 = 12.33$, $p < .01$ $\mathrm{Min}F'(1,38) = 8.96$, $p < .01$	$F_1 = 1.94$, $p = .17$ $F_2 = 2.18$, $p = .15$ $\mathrm{Min}F'(1,54) = 1.03$, $p = .32$	Fs < 1.5
Percent regressions	$F_1 = 10.54$, $p < .01$ $F_2 = 7.74$, $p < .02$ $\mathrm{Min}F'(1,49) = 4.46$ $p < .05$	$F_1 = 6.97$, $p < .02$ $F_2 = 7.37$, $p < .02$ $\mathrm{Min}F'(1,54) = 3.58$, $p = .06$	Fs < 1.5

Note. For all analyses by participants (F_1) , $df = (1,34)$; for all analyses by items (F_2), $df = (1,22)$.

difference between conditions on the percent regressions measure, and this resulted from more regressions in the obligatory condition than in the optional condition (11.4% vs. 6.5%).

The first pass results supported our hypothesis that the shifted noun phrase would induce relative processing difficulty when the verb was optionally transitive. However, the fact that there were more regressions from the noun phrase region in the obligatory condition than in the optional condition (leading to a numerically greater go-past time in the obligatory condition) was not predicted. We think that the overall results support the notion that the noun phrase region was easier to process when the verb was obligatorily transitive, since the percent regressions data reflect a very small proportion of trials, and the first pass data, by contrast, reflect processing on every trial. (For discussion of the interpretation of the first pass and regression measures, see Altmann, 1994; Altmann, Garnham, & Dennis, 1992; Rayner & Sereno, 1994a, 1994b.) Nevertheless, one of our goals for Experiment 2 was to clarify the pattern of results on the noun phrase region. To anticipate the results from that experiment, it appears that the difference in first pass time is replicable, but the difference in percent regressions is not.

Spillover region

There were no differences that approached significance on this region, on any measure. Evidently, all effects of the experimental manipulation were completely resolved by the time readers' eyes moved past the head noun of the shifted noun phrase.

In sum, the results of Experiment 1 indicated quite clearly that a post-verbal prepositional phrase causes processing difficulty when the verb is obligatorily transitive, compared to when the verb is transitive-biased but can occur in an intransitive frame. In addition, the first pass results supported our prediction of processing difficulty on the shifted noun phrase itself when the verb was optionally transitive compared to when it was obligatorily transitive. To perform a direct test of our prediction of a crossover pattern, with the heavy NP shift word order causing difficulty on the prepositional phrase in the obligatory condition and difficulty on the noun phrase in the optional condition, we performed a 2×2 ANOVA on first pass time, with verb type and region as factors. The main effect of verb type did not approach significance, but the interaction of verb type and region was highly significant: $F_1(1,34) = 17.81$, $p < .01$; $F_2(1,22) = 8.41$, $p < .01$; $\mathrm{Min}F'$ $(1,41) = 5.71$, $p < .05$. We reserve further discussion of these results until we have presented the results of Experiment 2.

Experiment 2

Experiment 2 had several purposes. First, we wanted to determine whether the results of Experiment 1 would generalize to heavy NP shift sentences in which different material intervened between the verb and the shifted direct object. In Experiment 2, we used an adverb rather than a prepositional phrase as this intervening material.

Second, we wanted to compare the processing of heavy NP shift sentences with obligatorily and optionally transitive verbs with the processing of unshifted sentences with the same verbs. We have interpreted the evidence from Experiment 1 of relative processing difficulty on the prepositional phrase when the verb is obligatorily transitive, and relative processing difficulty on the shifted noun phrase when the verb is optionally transitive, as confirming our proposal that the parser constructs the heavy NP shift analysis at one of the these points, depending on verb type. Our account predicts, therefore, that with an unshifted word order (i.e., when

the direct object appears immediately after the verb), the effects we have observed should disappear.

Third, we suspected that the percent regressions difference on the noun phrase region may not have been a reliable effect, primarily because it contradicted the first pass results. We thought that a replication of Experiment 1 might confirm the first pass effect on the noun phrase region without replicating the percent regressions effect.

Finally, we were interested in whether transitivity preferences play a major role in determining when the parser adopts the heavy NP shift analysis. In Experiment 1, all of the optionally transitive verbs that we used had a high transitive bias. In Experiment 2, we used optionally transitive verbs whose biases ranged more widely, and had only a moderate transitive bias overall. This would enable us to investigate, in a different way, the prediction of the lexical expectation strategy that the likelihood of the parser adopting the heavy NP shift analysis before reaching the shifted noun phrase is correlated with the verb's transitivity bias.

In Experiment 2, we tested sentences like the following:

(7a) The teacher immediately corrected the unusual answer the student had given.
(7b) The teacher immediately applauded the unusual answer the student had given.
(7c) The teacher corrected immediately the unusual answer the student had given.
(7d) The teacher applauded immediately the unusual answer the student had given.

In (7a) and (7c), the verb is obligatorily transitive, while in (7b) and (7d), the verb is optionally transitive. In (7a) and (7b), the adverb (*immediately*) appears before the verb, while in (7c) and (7d), the adverb appears after the verb, with the verb's direct object (*the unusual answer the student had given*) shifted to the right. Accordingly, we refer to (7a–d) respectively, as the *adverb first obligatorily transitive* condition, the *adverb first optionally transitive* condition, the *adverb second obligatorily transitive* condition and the *adverb second optionally transitive* condition.

Sentences (7c) and (7d) are analogous to the obligatory and optional conditions in Experiment 1, and we predicted the same pattern of results: relative difficulty on the material intervening between verb and noun phrase in (7c), and relative difficulty on the shifted noun phrase in (7d). We did not expect the percent regressions effect on the noun phrase region that appeared in Experiment 1 to replicate in this experiment; essentially, we suspected that this finding was simply a Type I error. If it is in fact the case that the appearance of the adverb after the verb in (7c) causes disruption, and causes the parser to construct the heavy NP shift analysis, then processing the adverb should also be more difficult in (7c) than in (7a), in which the adverb comes before the verb, and there should be no difference between these sentences on the noun phrase region. On the other hand, processing the direct object noun phrase should be more difficult in (7d) than in (7b), with no difference between these conditions on the adverb.

Method

Participants

Twenty-four native speakers of American English, who were students at the University of Massachusetts, were given course credit or were paid $5 to participate in the experiment. All had normal or corrected-to-normal vision, and all were naïve to the purpose of the experiment.

Materials

We constructed 28 sets of sentences similar to (7a–d) above, in which verb type and adverb position were manipulated. In the (a) and (c) versions of each sentence, we (and several other raters) judged a direct object noun phrase to be required after the main verb in order for the sentence to be grammatical. In the (b) and (d) versions, on the other hand, the main verb allowed the sentence to continue either with or without a direct object. For example, it is possible to construct a version of (7d) in which the adverb is followed by an adjunct clause: *The teacher applauded immediately when the student gave her unusual answer*. The full set of materials is provided in Appendix B.

The 28 sets of sentences used 14 different verb pairs, with each verb pair appearing in two sets of sentences. All but one of the verb pairs was matched for length, with the verbs in the remaining pair differing in length by one character. The obligatorily and optionally transitive verbs did not differ in frequency, based on the Francis and Kučera (1982) norms. The obligatorily transitive verbs had a mean $\log_n$ frequency of 3.84, $SD = 1.50$, and the optionally transitive verbs had a mean $\log_n$ frequency of 4.13, $SD = 1.65$. Four lists were constructed so that each participant saw one sentence from each set, and seven sentences overall in each of the four conditions. In addition, as in Experiment 1 the lists of stimuli were arranged so that each participant saw each verb only once; i.e., if a participant saw a version of (7) that included the verb *corrected*, he or she would see a version of the other *corrected/applauded* stimulus set that included the verb *applauded*.

The 14 optionally transitive verbs that we used varied widely in their transitivity biases. As in Experiment 1, we obtained transitivity preferences for these verbs from the Connine et al. (1984) and Gahl et al. (2004) norms. Twelve of the verbs were included in the Connine et al. norms, and 13 were included in the Gahl et al.

norms. We calculated the transitivity bias of the 12 verbs that appeared in both sources as the mean of the scores from these two sources; the correlation between the scores from the two sources was $r = .44$. For *report* we used only the Gahl et al. norms, since this verb did not appear in the Connine et al. norms. For *applaud*, which appeared in neither set of norms (and which happens to be the optionally transitive verb in our sample set of sentences above), we conducted a production norming study with 10 participants, and determined the transitivity bias of this verb to be .60. Table 4 provides the transitivity biases obtained from each set of norms, as well as the means of these scores. Overall, the transitivity biases of the optionally transitive verbs ranged from .04 (*try*) to .82 (*watch*), with a mean of .50 and $SD = .22$.

The optionally transitive verbs that we used vary with respect to whether they frequently occur with an argument other than a noun phrase. Some of these verbs (e.g., *cheat*, *walk*) frequently occur as pure intransitives, i.e., with no internal argument at all. Others, however, do tend to take an argument other than a noun phrase. For example, *try* tends to occur with a non-finite clausal complement (e.g., *try to fix the car*), and *report* tends to occur with a finite clausal complement (e.g., *report that the battle was over*). These differences are glossed over in computing overall transitivity preference, both in our own experiment and, frequently, in others in which verb transitivity bias is studied (see Gahl et al., 2004, for discussion of related issues). We acknowledge that these differences may be relevant to the comprehension of heavy NP shift; Stallings et al. (1998) present evidence that they play a role in the production of heavy NP shift. We briefly revisit this issue in the General discussion.

Table 4
Experiment 2: Transitivity biases of optionally transitive verbs

Verb	Connine et al. (1984)	Gahl et al. (2004)	Overall
worship	.72	.60	.66
perform	.32	.76	.54
report	—	.34	.34
follow	.94	.66	.80
pull	.16	.57	.37
watch	.90	.73	.82
try	.03	.05	.04
applaud	—	—	—
study	.32	.86	.59
cheat	.22	.55	.39
lecture	.28	.38	.33
protest	.62	.25	.44
help	.80	.80	.80
debate	.20	.49	.35

Note. The transitivity bias of *applaud*, which appeared in neither set of norms, was assessed by conducting a single-item production norming session with ten participants. We obtained a transitivity bias of .60.

Of the 14 obligatorily transitive verbs used in this experiment, seven were repeated from Experiment 1. Because four of the remaining seven verbs appeared in neither the Connine et al. nor the Gahl et al. norms, we had 16 participants construct sentences using these seven verbs, which were intermixed with optionally transitive and intransitive verbs. One response was excluded because it used the critical word as a noun, leaving a total of 111 responses. Of these, three responses were clearly idiomatic (e.g., *Buy low, sell high*), and five made use of a particle verb construction (e.g., *Joe blocked off the driveway*). The latter are clearly not relevant to the present experiment, since the particle in such constructions can only be separated from the verb by a direct object noun phrase. Of the remaining 103 responses, 102 included a noun phrase direct object for the verb (99%).

The experimental sentences were intermixed with 80 filler sentences of various types, none of which involved the heavy NP shift word order. The sentences were presented in an individually randomized order to each participant.

Procedure

The procedure was identical to Experiment 1. Average accuracy on the comprehension questions was above 90%, with no participant scoring below 80%. The entire experiment lasted approximately 30 min.

Results and discussion

We analyzed three regions in each sentence that corresponded to the three regions we analyzed in Experiment 1. The *adverb region* consisted only of the adverb (*immediately*), which appeared before the verb in the adverb first conditions, and after the verb in the adverb second conditions. The next region, the *noun phrase region*, consisted of the first three words of the direct object noun phrase, which included a determiner, an adjective, and the head noun (*the unusual answer*). Finally, we analyzed a *spillover region* that consisted of the next two words of the sentence (*the student*).

We computed the same four reading time measures as in Experiment 1: *first fixation duration*, *first pass time*, *go-past time*, and *percent regressions*. We point out that as in Experiment 1, the first fixation duration measure is meaningful only for the first two of the three regions, where the beginning of the region marked a phrasal boundary. Less than 2% of all trials were excluded due to track losses. We used the same procedure as in Experiment 1 to combine very short fixations with neighboring fixations. Less than 1.5% of remaining fixations were eliminated due to falling outside the 80–800 ms range.

For each measure on each region, we first performed ANOVAs by participants and by items, with condition as single within-participants or within-items factor with

four levels, and with counterbalancing group as a between-participants or between-items factor. To test our hypotheses in the most direct manner possible, we then tested three specific contrasts: adverb first obligatorily transitive versus adverb second obligatorily transitive; adverb first optionally transitive versus adverb second optionally transitive; and adverb second obligatorily transitive versus adverb second optionally transitive. Table 5 presents the condition means, as well as the confidence intervals for each of the critical differences between conditions. Table 6 presents the F statistics for the omnibus ANOVAs. We report a difference as significant if both subjects and items analyses rejected the null hypothesis at the .05 level. We discuss the regions individually below.

Adverb region

Go-past time on the adverb in the adverb second obligatorily transitive condition (449 ms) was significantly longer than in the adverb second optionally

Table 5
Experiment 2: Participant mean reading times, in ms, and percent regressions

Measure	Adverb region (*immediately*)	NP region (*the unusual answer*)	Spillover (*the student*)
First fixation duration			
Adverb first/obligatory	287	259	263
Adverb first/optional	288	264	262
Adverb second/obligatory	294	267	251
Adverb second/optional	289	282	252
95% CI of difference between two means	±15.8	±19.5	±18.9
Contrast 1	−22.0 to 9.6	−27.5 to 11.5	−6.6 to 31.2
Contrast 2	−17.0 to 14.6	−37.8 to 1.2	−8.5 to 29.3
Contrast 3	−11.2 to 20.4	−34.6 to 4.4	−17.4 to 20.4
First pass time			
Adverb first/obligatory	348	746	383
Adverb first/optional	335	752	369
Adverb second/obligatory	358	756	356
Adverb second/optional	344	847	356
95% CI of difference between two means	±25.5	±61.3	±29.3
Contrast 1	−35.4 to 15.6	−70.7 to 51.9	−2.3 to 56.3
Contrast 2	−33.6 to 17.4	−156.1 to −33.5*	−16.6 to 42.0
Contrast 3	−11.1 to 39.9	−152.4 to −29.8*	−29.5 to 29.1
Go-past time			
Adverb first/obligatory	391	868	403
Adverb first/optional	380	798	406
Adverb second/obligatory	449	946	374
Adverb second/optional	364	969	430
95% CI of difference between two means	±38.9	±72.8	±51.6
Contrast 1	−96.6 to −18.8*	−151.0 to −5.4	−22.4 to 80.8
Contrast 2	−22.8 to 55.0	−244.1 to −98.5*	−75.6 to 27.6
Contrast 3	46.4 to 124.2*	−95.6 to 50.0	−107.6 to −4.4
Percent regressions			
Adverb first/obligatory	6.1	10.8	2.8
Adverb first/optional	6.6	5.8	7.8
Adverb second/obligatory	15.6	13.7	2.8
Adverb second/optional	4.3	11.3	7.3
95% CI of difference between two means	±5.7	±6.4	±5.7
Contrast 1	−15.2 to −3.8*	−9.3 to 3.5	−5.7 to 5.7
Contrast 2	−3.4 to 8.0	−11.9 to 0.9	−5.2 to 6.2
Contrast 3	5.6 to 17.0*	−4.0 to 8.8	−10.2 to 1.2

Adverb first/obligatory = *The teacher immediately corrected the unusual answer the student had given.* Adverb first/optional = *The teacher immediately applauded the...* Adverb second/obligatory = *The teacher corrected immediately the...* Adverb second/optional = *The teacher applauded immediately the...* Contrast 1: *Adverb first obligatory vs. adverb second obligatory.* Contrast 2: *Adverb first optional vs. adverb second optional.* Contrast 3: *Adverb second obligatory vs. adverb second optional.*

Note. Reported here are the 95% confidence intervals for the differences between conditions, calculated using the procedure for within-participant confidence intervals described by Loftus and Masson (1994; Masson and Loftus, 2003). Contrasts that were significant by both participants and items ($p < .05$) are marked with an asterisk (*).

Table 6
Experiment 2: One-way analysis of variance results by region for effect of condition

Measure	Adverb	Noun Phrase	Spillover
First fixation duration	Fs < 1.5	$F_1 = 2.06$, $p = .12$ $F_2 = 2.48$, $p = .07$ Min$F'(3,127) = 1.13$, $p = .34$	Fs < 1.5
First pass time	Fs < 1.5	$F_1 = 4.88$, $p < .01$ $F_2 = 4.81$, $p < .01$ Min$F'(3,131) = 5.13$, $p = .07$	$F_1 = 1.54$, $p = .21$ $F_2 = 1.03$, $p = .39$ Min$F'(3,130) = .62$, $p = .61$
Go-past time	$F_1 = 7.27$, $p < .01$ $F_2 = 4.34$, $p < .01$ Min$F'(3,128) = 2.72$, $p < .05$	$F_1 = 9.19$, $p < .01$ $F_2 = 6.69$, $p < .01$ Min$F'(3,131) = 3.87$, $p < .02$	$F_1 = 1.58$, $p = .20$ $F_2 = .49$, $p = .69$ Min$F'(3,110) = .37$, $p = .77$
Percent regressions	$F_1 = 6.38$, $p < .01$ $F_2 = 5.42$, $p < .01$ Min$F'(3,131) = 2.93$, $p < .05$	$F_1 = 2.11$, $p = .11$ $F_2 = 1.65$, $p = .19$ Min$F'(3,131) = .93$, $p = .43$	$F_1 = 1.89$, $p = .14$ $F_2 = 1.47$, $p = .23$ Min$F'(3,131) = .83$, $p = .48$

Note. For all analyses by participants (F_1) , $df = (3,60)$; for all analyses by items (F_2), $df = (3,72)$.

transitive condition (364 ms) and the adverb first obligatorily transitive condition (391 ms). There were also significantly more regressions from the adverb in the adverb second obligatorily transitive condition (15.6%) than in the adverb second optionally transitive condition (4.3%) and the adverb first obligatorily transitive condition (6.1%). The adverb second obligatorily transitive condition did not differ from the other conditions on the first fixation or first pass measures. The two optionally transitive conditions did not differ significantly on any measure.

In sum, there was clear evidence that the adverb induced processing difficulty when it followed an obligatorily transitive verb, compared to when it followed an optionally transitive verb or when it preceded an obligatorily transitive verb, in the form of more regressive eye movements and longer go-past reading times. However, this experiment did not significantly replicate the first pass effect that appeared on the corresponding prepositional phrase region in Experiment 1. We attribute this difference simply to the fact that the adverbs were much shorter than the prepositional phrases, allowing less time for a first pass effect to appear.

Noun phrase region

On the noun phrase region, readers' first fixation duration in the adverb second optionally transitive condition (282 ms) was marginally longer than in the adverb second obligatorily transitive condition (267 ms) and the adverb first optionally transitive condition (264 ms). However, first pass time in the adverb second optionally transitive condition (847 ms) was significantly longer than in both the adverb second obligatorily transitive condition (756 ms) and the adverb first optionally transitive condition (752 ms). We note that the absolute size of the first pass effect on the noun phrase region was more than twice as great as in Experiment 1.

Go-past time was significantly longer in the adverb second optionally transitive condition (969 ms) than in the adverb first optionally transitive condition (798 ms). However, the difference between the adverb second optionally transitive condition and the adverb second obligatorily transitive condition (946 ms) did not approach significance. Finally, the greatest number of regressions occurred in the adverb second obligatorily transitive condition, as in the analogous condition of Experiment 1. However, in contrast to Experiment 1, the difference between the adverb second obligatorily transitive condition (13.7%) and the adverb second optionally transitive condition (11.3%) did not approach significance. In addition, we note that the shifted word order resulted in at least as large an increase in the probability of regressing from the noun phrase region when the verb was optionally transitive (5.8–11.3%) as when the verb was obligatorily transitive (10.8–13.7%). Clearly, the finding of Experiment 1 of an increase in the probability of regressing following an obligatorily transitive verb is not associated with the processing of heavy NP shift.

The two obligatorily transitive conditions did not differ significantly on any measure. There was a numerical difference in go-past time that was significant in the participants analysis, but not in the items analysis ($p = .12$).

Spillover region

The paired comparisons revealed no significant differences between conditions, on any measure. (The two adverb second conditions differed significantly in go-past time in participants analysis, but this difference did not approach significance in the items analysis.) However, a post hoc 2×2 ANOVA with word order and verb type as factors showed a significant effect of verb type on the percent regressions measure, with more first pass regressions from this region when the main verb was optionally

transitive than when it was obligatorily transitive: $F_1(1,20) = 5.09, \quad p < .05; \quad F_2(1,24) = 5.02, \quad p < .05;$ $\mathrm{Min}F'(1,43) = 2.53, p = .12$. We note, however, that even in the optionally transitive conditions there were relatively few regressions, and the influence of verb type on regressions from the spillover region did not manifest itself in a difference in go-past time. In general, it appears that readers resolved whatever difficulty was caused by the heavy NP shift word order by the time their eyes moved past the head noun of the shifted noun phrase.

The overall pattern of results is easily summarized. The first prediction was that the adverb would induce processing difficulty only when it appeared following an obligatorily transitive verb. This was confirmed by the go-past and percent regressions data. Readers made more regressive eye movements from the adverb when it followed an obligatorily transitive verb than when it preceded the verb or when it followed an optionally transitive verb, and they took longer to move past the adverb to the right. The second prediction was that the direct object noun phrase would induce processing difficulty with the shifted word order, but only when the verb was optionally transitive. Reading time was increased on the first three words of the direct object noun phrase in this condition compared to when the sentence appeared in an unshifted order and compared to when the verb was obligatorily transitive, beginning with the first fixation on this region and reaching full significance in first pass reading time. This effect did not, however, manifest itself in an increase in regressive eye movements or go-past time. Finally, there were essentially no interpretable effects of the experimental manipulation on the spillover region following the head noun of the shifted noun phrase.

It is important to note that there was no hint of difficulty on the post-verbal adverb when the verb was optionally transitive, and there were only very slight (and non-significant) signs of difficulty on the shifted noun phrase when the verb was obligatorily transitive. We regard these negative findings as further confirmation of our hypothesis about how the reader is likely to proceed through a heavy NP shift construction. Evidently, readers found it quite easy to adopt an intransitive frame when they encountered an adverb after a verb that allowed such a frame. When forced to adopt the heavy NP shift analysis at this point, because the verb was obligatorily transitive, readers had trouble doing so, but they were able to do so by the time they reached the direct object itself, so that this direct object was easily integrated.

It is worth noting, and ruling out, a possible objection to the design of this experiment. In the adverb second obligatorily transitive condition, the reader does in fact have available a grammatical analysis other than heavy NP shift, upon reaching the adverb. This alternate analysis initially escaped our notice. Consider (8):

(8) The teacher corrected immediately replying students, but not those who waited to give an answer.

In this sentence, the adverb *immediately* is part of an adjective phrase that is itself part of the noun phrase that is the verb's direct object. If participants adopted this analysis upon reading the adverb, it would seriously complicate the interpretation of the experimental results; fortunately, there is good reason to doubt that they did so. First, the adverb in the adverb second obligatorily transitive condition was at least as disruptive as the post-verbal prepositional phrase in the obligatory condition of Experiment 1, as measured by percent regressions and go-past time. Second, Hagoort, Brown, and Groothusen (1993) found in an ERP experiment with Dutch speakers that an adverb behaved like a syntactic anomaly in a position in which it could only legally modify a subsequent adjective, evoking a clear "syntactic positive shift." In other words, though this syntactic analysis is a legal one, readers do not seem to recognize this analysis on-line without clear evidence favoring it.

As discussed above, we hoped that Experiment 2 would help to address the question of whether transitivity bias plays a major role in the processing of heavy NP shift. To answer this question, we computed two difference scores for each set of sentences. The first of these difference scores, which we will call the adverb difference score, was simply the go-past time on the adverb region in the adverb second optionally transitive condition minus the go-past time on the adverb region in the adverb first optionally transitive condition. The lexical expectation account would predict that across our 28 items, there should be a positive correlation between an item's adverb difference score and the transitivity bias of the optionally transitive verb that appeared in that item, since this account predicts that the parser will construct the heavy NP shift structure when an adverb follows a verb that is preferentially transitive. The second difference score, which we will call the noun phrase difference score, was the first pass time on the noun phrase region in the adverb second optionally transitive condition minus the first pass time on the noun phrase region in the adverb first optionally transitive condition. The lexical expectation account would predict a *negative* correlation between an item's noun phrase difference score and the transitivity bias of the optionally transitive verb, since this account predicts that when a verb has a high transitive bias the heavy NP shift analysis is likely to be adopted earlier in the sentence, and as a result, a shifted direct object should be more easily processed.

The results of these analyses do not support the notion that transitivity bias determines the probability of an early adoption of the heavy NP shift analysis. Both correlations are remarkably close to zero. The correlation between the adverb difference scores and

the transitivity bias of the optionally transitive verbs was $r = .002$, $p = .99$; the correlation between the noun phrase difference scores and these transitivity biases was $r = .023$, $p = .91$. Though this is merely negative evidence, we think it is difficult to explain on an account of the processing of heavy NP shift that emphasizes the role of transitivity bias.

General discussion

The two experiments presented in this paper paint a very coherent picture of the processing of heavy NP shift constructions. Experiment 2 shows clearly that the heavy NP shift word order causes processing difficulty, compared to the canonical word order in which the direct object appears immediately after the verb. However, both Experiments 1 and 2 show that the point at which this difficulty arises varies depending on the verb's subcategorization possibilities: when the verb is obligatorily transitive, difficulty appears on the material that intervenes between the verb and its direct object, but when the verb is optionally transitive, difficulty does not appear until the reader reaches the direct object itself. Experiment 1 shows that this difference exists even with optionally transitive verbs that are strongly transitive-biased; Experiment 2 suggests that transitivity bias, within the set of optionally transitive verbs, may matter relatively little.

One question about the results of these experiments is why the effects of the heavy NP shift word order on the intervening material between the verb and its direct object (when the verb was obligatorily transitive) and the shifted noun phrase (when the verb was optionally transitive) were manifested in different eye movement measures. When readers encountered a prepositional phrase (in Experiment 1) or an adverb (in Experiment 2) following an obligatorily transitive verb, they were relatively likely to make regressive eye movements back to an earlier region of the sentence, resulting in longer go-past times. In Experiment 1, there was also a difference in first pass reading time on the prepositional phrase, though this was not significant on the adverb in Experiment 2. On the other hand, when readers encountered a direct object noun phrase following an optionally transitive verb and an adverb or prepositional phrase, they did not make an especially large number of regressive eye movements, but instead they merely slowed down while continuing to make forward saccades, in both experiments.

We think that some light can be shed on this pattern by considering the nature of the difficulty that the parser encounters at each point. When an adverb or prepositional phrase appears immediately after an obligatorily transitive verb, there is no way for this constituent to be attached directly into the phrase marker. We think that readers may regress from this region because they are initially at a loss about how to repair or revise the structure they have built; perhaps they are checking to make sure that they have not misread the verb. Ultimately, they must posit a trace of movement in the direct object position, but there is no obvious clue to the reader that this is what is required. On the other hand, when an optionally transitive verb is followed by an adverb or prepositional phrase, and then by a noun phrase, this noun phrase itself provides evidence that what is required is a heavy NP shift structure. As a result, the reader need not make regressive eye movements in an attempt to diagnose the problem and to discover a possible solution.

The experiments we present here point toward several theoretically interesting conclusions. First, it is clear that in the processing of heavy NP shift the parsing system honors grammatical requirements as they arise. Evidently, the parser incrementally builds a highly specified phrase marker to which each new phrase must be attached in a manner that is consistent with the verb's subcategorization requirements. When a post-verbal prepositional phrase or adverb can only be attached by constructing the heavy NP shift analysis, there is evidence that the parser does indeed construct this alternate analysis rather than blithely proceeding. At the same time, the evidence of difficulty on the shifted noun phrase with optionally transitive verbs suggests that the parser becomes committed to an intransitive frame rather early in the sentence, instead of leaving its options open. In principle, the parser could adopt a delay strategy (Marcus, 1980), selecting a specific verb frame only after the heavy NP shift analysis has been ruled out. Alternately, the parser could construct a partially specified or "good enough" representation (Christianson, Hollingworth, Halliwell, & Ferreira, 2001), in which it is left indeterminate whether or not the verb has a direct object. In fact the parser does not take either of these options.

These experiments also reiterate a well-established methodological point: eye movements during reading reveal fine details of language processing that are not introspectively apparent. In our experience, people are sometimes aware that heavy NP shift constructions are at least somewhat difficult to process, but they certainly are not able to identify the specific points at which this difficulty arises, or how the nature of the difficulty depends on verb type.

The experiments clearly support a last resort strategy for the processing of heavy NP shift, at least in contrast to the specific alternatives that we outlined in the introduction. (We mention other alternatives below.) We motivated the hypothesis that the parser would adopt a last resort strategy by appealing to the preference for immediate incremental interpretation: if the parser assumes an intransitive frame when it encounters

something other than a direct object noun phrase after a verb, the semantic processor can immediately begin interpreting this material as a modifier of the verb. However, we note that it is possible to explain the last resort strategy by appealing to a structural preference embodied in De Vincenzi's (1991) Minimal Chain Principle. This principle says that the parser posits necessary chain members (traces) of moved constituents as soon as possible, but avoids positing unnecessary chain members. Consequently, it predicts the first resort strategy for the processing of *wh*-movement, and the last resort strategy for the processing of heavy NP shift. We suspect, however, that at some level these two explanations are not truly distinct. De Vincenzi explains the existence of the Minimal Chain principle by pointing out that adopting this principle enables the language comprehender to rapidly structure incoming material, and that "obviously the ultimate reason to structure materials is to reach the semantic interpretation of each individual sentence... [T]he parser is under constant pressure to pass each constituent to a semantic interpreter as soon as possible" (1991, p. 148).

Finally, the experiments presented here suggest that the transitivity bias of an optionally transitive verb that appears in a heavy NP shift construction is not a very important factor in determining how this construction is processed. In this respect, our results conflict with previous studies that have found transitivity biases to play an important role in initial parsing preferences (e.g., Garnsey, Perlmutter, Myers, & Lotocky, 1997; Trueswell, Tanenhaus, & Kello, 1993) and are consistent with other studies finding a delayed effect of verb biases, if an effect appeared at all (e.g., Ferreira & Henderson, 1990; Kennison, 2001; Pickering & Traxler, 2003; Pickering, Traxler, & Crocker, 2000). We wish to emphasize that the results of the present experiments do not suggest that frequency plays no role in initial parsing preferences. What they do suggest is that when frequency information conflicts with the preference for rapidly obtaining an interpretable grammatical analysis, the latter is likely to win out.

We note, furthermore, that there are at least two frequency-based accounts that seem consistent with the results of the present experiments. First, it is possible that what is responsible for the parser's reluctance to adopt the heavy NP shift analysis, until it has no other option, is the relatively low overall frequency of the heavy NP shift word order. We noted in the introduction that Wasow (1997) found, in a study of printed materials, that 5–10% of sentences with direct objects used the heavy NP shift word order. It is possible that the low frequency of heavy NP shift overrides the frequency with which a specific verb appears with a direct object. While a specific verb may be quite likely to appear in a transitive frame, it is probably true that in general a verb followed by an adverb or prepositional phrase is more likely to signal an intransitive verb phrase than heavy NP shift. Perhaps it is the relative frequency of these syntactic constructions in general, rather than verb-specific transitivity biases, that determines parsing preferences. Second, it is possible that what matters is the verb-specific frequency of the heavy NP shift word order rather than the verb's transitivity bias. Stallings et al. (1998) have shown that some of the variation in how frequently specific verbs appear with the heavy NP shift word order is systematic variation due to argument structure differences. It is possible that the parser is quick to adopt the heavy NP shift analysis with precisely those verbs that tend to be produced with the heavy NP shift word order. We note that there is evidence against at least one version of this account. As discussed above, Wasow (1997) found that heavy NP shift was more common with optionally transitive verbs than with obligatorily transitive verbs; but the experiments presented in this article demonstrate that it is specifically with optionally transitive verbs that the parser delays adopting the heavy NP shift analysis until the shifted noun phrase is encountered. Nevertheless, we think it is plausible that the low overall frequency of heavy NP shift and/or verb-specific variation in heavy NP shift frequency play some role in the processing of this construction. Obviously, more experimentation is required to settle these questions.

Appendix A. Experiment 1—Items

For each item, the verb that differed between versions is set off by slashes. The first verb is the obligatorily transitive version. The regions used for analysis are described in the text.

1. The sailors persuaded/worshiped in the morning the old pirate who had come aboard.
2. Billy persuaded/worshiped with great enthusiasm the cool kid with the red bike.
3. The popular king included/governed for many years the poor peasants in the nearby towns.
4. The general included/governed on the battlefield all the troops under his command.
5. Sarah bothered/attacked with no mercy the family of ants living on the windowsill.
6. Richard bothered/attacked at twelve o'clock the red-haired kid who lived on his street.
7. The math teacher praised/watched during the activity the quiet students in the back row.
8. Jack praised/watched from the stands his daughter's attempt to shoot a basket.
9. The professor sold/left in the winter the old sports car that had sentimental value.
10. Karen sold/left with great regret the last remaining copy of her first novel.
11. Roger invited/visited on many evenings all the other faculty in the department.

12. The plant manager invited/visited whenever possible the majority owners of the company.
13. Sue imitated/followed with little difficulty the advanced tricks she saw at the circus.
14. The kitten imitated/followed in the backyard the shaggy dog who could catch a ball.
15. The gardener placed/pushed against the shed a full wagon-load of flowering plants.
16. Kevin placed/pushed with great care the tiny metal soldiers that came from France.
17. Rachel guarded/painted for the whole night the watercolor landscape on the easel.
18. The wine maker guarded/painted in the cellar the special racks for the new wine.
19. The committee hired/heard over the weekend the visiting lecturer from Chicago.
20. Tim hired/heard after the meeting a new manager who was an expert on risk assessment.
21. George coaxed/taught almost every day the troubled students who had little enthusiasm.
22. The trainer coaxed/taught by himself the frisky puppy who wouldn't come inside.
23. The bellboy carried/cleaned for most of the day the convention guests' many suitcases.
24. Joe carried/cleaned with some reluctance the disgusting garbage on the floor.

Appendix B. Experiment 2—Items

For each item, the material that differed between versions is set off by slashes. The first verb is the obligatorily transitive version. The regions used for analysis are described in the text.

1. The crowd publicly destroyed/publicly worshiped/destroyed publicly/worshiped publicly the holy statue made of bronze.
2. The visitors voluntarily destroyed/voluntarily worshiped/destroyed voluntarily/worshiped voluntarily the gigantic sculpture with two heads.
3. The actor briefly described/briefly performed/described briefly/performed briefly the hilarious monologue from the play.
4. The villager eagerly described/eagerly performed/described eagerly/performed eagerly the ritual dance that he learned as a child.
5. The captain accurately mimicked/accurately reported/mimicked accurately/reported accurately the ridiculous behavior of the crew.
6. The manager angrily mimicked/angrily reported/mimicked angrily/reported angrily the juvenile antics of his players.
7. The pianist carelessly bothered/carelessly followed/bothered carelessly/followed carelessly the tempermental conductor with the wavy hair.
8. The teenager intentionally bothered/intentionally followed/bothered intentionally/followed intentionally the violent policeman in the blue uniform.
9. The farmer reluctantly bought/reluctantly walked/bought reluctantly/walked reluctantly the tired donkey with the sad eyes.
10. The trainer regretfully bought/regretfully walked/bought regretfully/walked regretfully the underfed greyhound who refused to race.
11. The producer frequently invited/frequently watched/invited frequently/watched frequently the experienced performers from Germany.
12. The professor occasionally invited/occasionally watched/invited occasionally/watched occasionally a guest lecturer from Mexico.
13. The store owner reluctantly hired/reluctantly tried/hired reluctantly/tried reluctantly the substitute barber with the weird hairdo.
14. The businessman voluntarily hired/voluntarily tried/hired voluntarily/tried voluntarily the personal trainer from the downtown gym.
15. The chef quietly corrected/quietly applauded/corrected quietly/applauded quietly the noble attempt made by the apprentice.
16. The teacher immediately corrected/immediately applauded/corrected immediately/applauded immediately the unusual answer the student had given.
17. The merchant nervously carried/nervously studied/carried nervously/studied nervously the expensive jewelry that came from India.
18. The worried child carefully carried/carefully studied/carried carefully/studied carefully the wounded rabbit with the soft fur.
19. The conqueror ruthlessly guarded/ruthlessly cheated/guarded ruthlessly/cheated ruthlessly the restless natives who lived on the island.
20. The old monkey cleverly guarded/cleverly cheated/guarded cleverly/cheated cleverly the younger members of the social group.
21. The teacher occasionally included/occasionally lectured/included occasionally/lectured occasionally the quiet students at the back of the class.
22. The warden primarily included/primarily lectured/included primarily/lectured primarily the newer inmates with bad attitudes.
23. The activist rudely imitated/rudely protested/imitated rudely/protested rudely the famous politician with the three-piece suit.
24. The performer comically imitated/comically protested/imitated comically/protested comically the annoying rules for entering the club.
25. The delivery boy repeatedly lifted/repeatedly helped/lifted repeatedly/helped repeatedly the handicapped child with the crutches.
26. The grandmother gladly lifted/gladly helped/lifted gladly/helped gladly the crying infant in the antique crib.
27. The officials regretfully blocked/regretfully debated/blocked regretfully/debated regretfully the final approval of the budget.
28. The janitors repeatedly blocked/repeatedly debated/blocked repeatedly/debated repeatedly the slippery spot in the fourth-floor hallway.

References

Aoshima, S., Phillips, C., & Weinberg, A. (2004). Processing filler-gap dependencies in a head-final language. *Journal of Memory and Language, 51*, 23–54.

Altmann, G. T. M. (1994). Regression-contingent analyses of eye movements during sentence processing: Reply to Rayner and Sereno. *Memory & Cognition, 22*, 286–290.

Altmann, G. T. M., Garnham, A., & Dennis, Y. (1992). Avoiding the garden path: Eye movements in context. *Journal of Memory and Language, 31*, 685–712.

Binder, K., Duffy, S., & Rayner, K. (2001). The effects of thematic fit and discourse context on syntactic ambiguity resolution. *Journal of Memory and Language, 44*, 297–324.

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*, 774–806.

Christianson, K., Hollingworth, A., Halliwell, J. F., & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive Psychology, 42*, 368–407.

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, The Netherlands: Kluwer.

Connine, C., Ferreira, F., Jones, C., Clifton, C., Jr., & Frazier, L. (1984). Verb frame preferences: Descriptive norms. *Journal of Psycholinguistic Research, 13*, 307–319.

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, UK: Cambridge University Press.

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

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

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

Ferreira, F., & Henderson, J. M. (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*, SI155–SI210.

Francis, W. N., & Kučera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston: Houghton Mifflin.

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

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII* (pp. 559–586). Hillsdale, NJ: Lawrence Earlbaum.

Frazier, L., & Clifton, C. Jr., (1989). Successive cyclicity in the grammar and the parser. *Language and Cognitive Processes, 28*, 331–344.

Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition, 6*, 291–326.

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

Gahl, S., Jurafsky, D., & Roland, D. (2004). Verb subcategorization frequencies: American English corpus data, methodolocial studies, and cross-corpus comparisons. *Behavior Research Methods, Instruments, and Computers, 36*, 432–443.

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

Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes, 8*, 439–483.

Kayne, R. (1998). Overt vs. covert movement. *Syntax, 1*, 128–191.

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

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

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review, 1*, 476–490.

Marcus, M. P. (1980). *A theory of syntactic recognition for natural language*. Cambridge, MA: MIT Press.

Masson, M. E. J., & Loftus, G. R. (2003). Using confidence intervals for graphically based data interpretation. *Canadian Journal of Experimental Psychology, 57*, 203–220.

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

Pickering, M. J., & Traxler, M. J. (2003). Evidence against the use of subcategorisation frequencies in the processing of unbounded dependencies. *Language and Cognitive Processes, 18*, 469–503.

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

Pollatsek, A., & Well, A. D. (1995). On the use of counterbalanced designs in cognitive research: A suggestion for a better and more powerful analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 785–794.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372–422.

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading*. Englewood Cliffs, NJ: Prentice Hall.

Rayner, K., & Sereno, S. C. (1994a). Regressive eye movements and sentence parsing: On the use of regression-contingent analyses. *Memory & Cognition, 22*, 281–285.

Rayner, K., & Sereno, S. C. (1994b). Regression-contingent analyses: A reply to Altmann. *Memory & Cognition, 22*, 291–292.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye-movement control in reading:

Comparisons to other models. *Behavioral and Brain Sciences, 26*, 445–526.

Rochemont, M., & Culicover, P. (1997). Deriving dependent right adjuncts in English. In D. Beerman, D. LeBlanc, & H. van Riemsdijk (Eds.), *Righward movement* (pp. 279–300). Amsterdam: John Benjamins.

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

Spivey, M. J., & 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.

Stallings, L. M., MacDonald, M. C., & O'Seaghdha, P. G. (1998). Phrasal ordering constraints in sentence production: Phrase length and verb disposition in heavy-NP shift. *Journal of Memory and Language, 39*, 392–417.

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

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.

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*, 528–533.

Wasow, T. (1997). End-weight from the speaker's perspective. *Journal of Psycholinguistic Research, 26*, 347–361.