The Sentence-Composition Effect: Processing of Complex Sentences Depends on the Configuration of Common Noun Phrases Versus Unusual Noun Phrases

Marcus L. Johnson, Matthew W. Lowder, and Peter C. Gordon
University of North Carolina at Chapel Hill

In 2 experiments, the authors used an eye tracking while reading methodology to examine how different configurations of common noun phrases versus unusual noun phrases (NPs) influenced the difference in processing difficulty between sentences containing object- and subject-extracted relative clauses. Results showed that processing difficulty was reduced when the head NP was unusual relative to the embedded NP, as manipulated by lexical frequency. When both NPs were common or both were unusual, results showed strong effects of both commonness and sentence structure, but no interaction. In contrast, when 1 NP was common and the other was unusual, results showed the critical interaction. These results provide evidence for a sentence-composition effect analogous to the list-composition effect that has been well documented in memory research, in which the pattern of recall for common versus unusual items is different, depending on whether items are studied in a *pure* or *mixed* list context. This work represents an important step in integrating the list-memory and sentence-processing literatures and provides additional support for the usefulness of studying complex sentence processing from the perspective of memory-based models.

Keywords: sentence complexity, memory, relative clauses, eye movements, list effects

Despite its rapid time course and seemingly effortless nature, sentence comprehension is a complex cognitive process that requires words to be encoded, stored, and retrieved from memory before they can be meaningfully integrated into a coherent interpretation. A large body of research on sentence processing has been devoted to gaining a better understanding of this process by looking at the contrast between the processing of object-extracted and subject-extracted relative clauses (RCs). In a subject-extracted RC, as in sentence 1A, the subject of the sentence also serves as the subject of the RC, whereas in an object-extracted RC, as in sentence 1B, the subject of the sentence serves as the object of the RC. A robust finding across many studies using different measures is that object RCs are substantially more difficult to process than subject RCs (e.g., Caplan, Alpert, & Waters, 1998; Caramazza & Zurif, 1976; Ford, 1983; Holmes & O'Regan, 1981; Just, Carpenter, Keller, Eddy, & Thulborn, 1996; King & Just, 1991; Wanner & Maratsos, 1978), a finding that can be attributed to sentence-interpretation processes because the two types of RCs contain exactly the same words and differ only in word order.

1A. The senator that bothered the reporter caused a big scandal.

1B. The senator that the reporter bothered caused a big scandal.

The basic object–subject asymmetry is virtually undisputed (but see Gennari & MacDonald, 2008), and it has been used as a test bed for exploring theories about the nature of sentence processing. In the present experiments, we investigated the effect of noun phrase (NP) frequency on the processing of RCs. In doing so, we were motivated by memory research showing that recall is strongly influenced by word frequency and that the dependence of this effect on the composition of a list of to-be-remembered items provides important insights into underlying mechanisms of memory. The results of the experiments show an analogous effect of the composition of a sentence in terms of the characteristics of its NPs; the pattern provides insight into the operations involved in sentence comprehension and shows that theoretical accounts of memory operation can deepen our understanding of language processing.

Memory-Based Explanations of the Object–Subject Processing Difference

Psycholinguistic models of the asymmetry in RC processing have placed a great deal of emphasis on how memory limitations affect the comprehension of complex sentences. Before reviewing those models, it is important to note that alternative approaches to this phenomenon have also been developed. One alternative class of theories highlights how semantic factors influence processing of

subject and object RCs, variously emphasizing the plausibility of the event described by the RC (King & Just, 1991), the semantic characteristics of the nouns in the RC (Mak, Vonk, & Schriefers, 2002, 2006; Traxler, Morris, & Seely, 2002; Traxler, Williams, Blozis, & Morris, 2005), or whether nouns in the RC have been mentioned recently in the discourse (Fox & Thompson, 1990; Gordon & Hendrick, 2005). A second alternative class of theories highlights the role of experience in explaining complexity effects on sentence processing, arguing that people have an easier time understanding types of sentences that they have encountered more frequently, as compared with those that they have encountered less frequently (Gennari & MacDonald, 2008; Hale, 2001; Levy, 2008). Semantic, experienced-based, and memory-based accounts of the asymmetry in processing subject and object RCs are not necessarily mutually exclusive; in principle, each could explain important aspects of this naturally occurring contrast. However, these accounts differ greatly in how they relate sentence processing to general principles of cognitive processing and in the predictions that those connections lead to with respect to factors that influence the ease of processing complex sentences.

Memory-based theories of complex sentence processing focus on the cognitive burden created during the comprehension of an object RC. Whereas subject RCs can be processed and understood incrementally, object RCs require the reader to hold two NPs in memory until they can be attached to verbs. Double-embedded object RCs, as in sentence 2A, have often been cited as an extreme example illustrating how a complex sentence structure can overburden memory and render accurate comprehension virtually impossible (Miller & Chomsky, 1963). Initial experimental research on the nature of this constraint made heavy use of memory-load tasks that examined how remembering a short list of words during comprehension affected the magnitude of the object–subject difference (Just & Carpenter, 1992; King & Just, 1991; Wanner & Maratsos, 1978; cf. Waters & Caplan, 1996, for a critique). More recent research on the role of memory in sentence processing has been heavily influenced by a phenomenon identified by Bever (1974), who noted that sentences that contain doubly, triply, and even quadruply embedded object RC structures become much easier to understand when the NPs of these structures are mixed in type rather than all descriptive terms (see the contrast between 2A and 2B).

2A.   The senator the banker the salesman knew trusts caused a big scandal.

2B.   The senator everyone I knew trusts caused a big scandal.

Two memory-based approaches have been advanced to account for how ease of RC processing is affected by the type of NPs used for the RC head (e.g., *the senator* in 1A & 1B) and the embedded NP (e.g., *the reporter* in 1A & 1B); for convenience, we refer to the head and embedded NPs as NP1 and NP2, respectively.

On one account, the resource demands imposed by representing NP2 in memory make it difficult to retrieve NP1 from memory. The dependence locality theory (Gibson, 1998, 2000; Grodner & Gibson, 2005; Warren & Gibson, 2002) proposes that a memory cost is incurred whenever an unattached constituent must be maintained in memory across intervening words, and an additional cost

is incurred when integration actually occurs. Critically, the costs that are incurred (i.e., the difficulty experienced by the reader) depend on the distance between the unattached constituent and its integration site. Object RCs (e.g., 1B) are difficult because NP1 must be maintained in memory across NP2 before it can be integrated with the embedded verb. In contrast, NP1 of a subject RC (e.g., 1A) can be immediately attached to the embedded verb. It is further argued (Warren & Gibson, 2002) that introducing a new discourse element in the sentence is more cognitively expensive to the reader than referring to a discourse element that is already cognitively available. This idea builds on the *givenness hierarchy* (Gundel, Hedberg, & Zacharski, 1993), in which referents that are more naturally given (e.g., "I," "you," "everyone") require fewer cognitive resources to access than referents that are not naturally given (e.g., "the senator," "the banker," "the salesman"). Warren and Gibson (2002) have presented reading-time data suggesting that the difficulty of processing object RCs is inversely related to the givenness of the intervening NPs.

An alternative approach stresses that the relationship between NP1 and NP2 can influence memory demands that contribute to the object–subject asymmetry in ease of processing. On this account, difficulty in comprehending object RCs occurs upon encountering the embedded verb because of interference in retrieval processes due to the presence of memory traces for two NPs; this interference is greater when NP1 and NP2 are similar, a pattern that supports the notion of similarity-based interference during memory retrieval (Gordon, Hendrick, & Johnson, 2001; Gordon, Hendrick, & Levine, 2002; see Lewis, 1999, for a different formulation of similarity-based interference). Empirical support for this idea has come from reading-time studies demonstrating that the object–subject difference in ease of processing is reduced or eliminated when the description used as NP2 (e.g., *the reporter*) is replaced by a proper name, indexical pronoun, or quantified pronoun, as shown in 3A and 3B. This reduction in the asymmetry appears with reading-time measures obtained using self-paced reading and eye tracking during normal reading (Gordon et al., 2001; Gordon, Hendrick, & Johnson, 2004; Gordon, Hendrick, Johnson, & Lee, 2006); the reduction is also seen in subjects' accuracy in answering postsentence comprehension questions.

3A.   The senator that [Bob/you/everyone] bothered caused a big scandal.

3B.   The senator that bothered [Bob/you/everyone] caused a big scandal.

Further empirical support for the idea that retrieval interference affects complex sentence processing comes from memory-load studies demonstrating that comprehension of sentences with object RCs is impaired when the words in the memory load are similar to the NPs in the complex sentence (Gordon et al., 2002; Van Dyke & McElree, 2006).

Research demonstrating similarity-based interference during the comprehension of complex sentences fits naturally within the cue-based parsing framework (Lewis & Vasishth, 2005; Lewis, Vasishth, & Van Dyke, 2006; Van Dyke & Lewis, 2003), which conceptualizes key aspects of sentence processing in terms of the memory processes of encoding, storage, and retrieval. Factors that can cause comprehension difficulty include interference from

other words in the sentence and long distances between the site of the encoded word and the point of retrieval. The emphasis in cue-based models on the processes of memory encoding and retrieval is consistent with the idea of *long-term working memory* (Ericsson & Kintsch, 1995), a construct that emphasizes rich encoding and effective retrieval strategies from long-term memory as a basis of skilled performance rather than emphasizing capacity limits of short-term memory as essential components of working memory (e.g., Baddeley, 1986, 2000). The ability to resume tasks with ease after interruption, which has been demonstrated for a number of tasks including language comprehension (Ericsson & Kintsch, 1995; Ledoux & Gordon, 2006), demonstrates the value of conceiving of working memory as relying on mechanisms of long-term memory. Below, we consider how manipulations and constructs developed by memory researchers studying effects on recall of the composition of lists of words may deepen understanding of memory encoding and retrieval operations during the processing of complex sentences.

Memory Encoding and Retrieval as a Function of List Composition

A well-documented finding in the memory literature is that patterns of free recall following the study of a list of words depends on list composition. For example, high-frequency words are better recalled than are low-frequency words after studying *pure* lists containing only high-frequency words or only low-frequency words. However, for *mixed* lists containing both high- and low-frequency words, free recall performance is typically greater for low-frequency words than high-frequency words (De-Losh & McDaniel, 1996; Duncan, 1974; Gregg, 1976; Gregg, Montgomery, & Castano, 1980; May & Tryk, 1970; Merritt, DeLosh, & McDaniel, 2006; cf. Hulme, Stuart, Brown, & Morin, 2003). This list-composition effect is not unique to the contrast between low- and high-frequency words; similar effects have also been observed for bizarre versus nonbizarre items (McDaniel, DeLosh, & Merritt, 2000; McDaniel, Einstein, DeLosh, May, & Brady, 1995), generated versus read items (Mulligan, 2002; Serra & Nairne, 1993), perceptually degraded versus intact items (Mulligan, 1999), humorous versus nonhumorous items (Schmidt, 1994), and orthographically irregular versus orthographically regular items (Hunt & Elliot, 1980). These studies support the idea that free-recall performance for a mixed list of common and unusual items will generally be better for the unusual items.

The item-order framework, originally proposed by Nairne, Riegler, and Serra (1991) and later developed by McDaniel and colleagues (DeLosh & McDaniel, 1996; McDaniel & Bugg, 2008), has been proposed as a theoretical explanation for the list-composition effect. According to this account, as we study a list of words, we not only encode item-specific information but also encode order-based information regarding each item's serial position on the list. When presented with a subsequent free-recall task, we rely heavily on this order-based information as a retrieval cue. Critically, this hypothesis also proposes that a list of low-frequency words tends to focus our attention more on the encoding of item-specific information, thus detracting from the encoding of order-based information. In contrast, studying a list of high-frequency words does not require as much item-focused attention, freeing up attention for the encoding of serial-position information,

thus making order information more available as a retrieval cue in pure lists of high-frequency words than in pure lists of low-frequency words. However, for mixed lists, the attention required to encode low-frequency words disrupts the encoding of order for the entire list, thereby reducing the value of a retrieval strategy based on order cues. With order-based information no longer serving as an effective retrieval cue, recall depends more on item-specific information, which is stronger for low- than for high-frequency words because of the greater effort expended to encode them. This explains why low-frequency words tend to be recalled better than high-frequency words in a mixed list.

The item-order account provides an attractive framework for considering how the characteristics of words in a sentence affect encoding and retrieval processes both for individual words and for the organization of the sentence. It draws on the same characterization of the basic mechanisms underlying successful memory performance that cue-based parsing approaches employ, describing memory performance as depending on the extent and the manner with which information is encoded and the degree to which the cues used in subsequent retrieval uniquely match the encoded features of the to-be-retrieved information. Although most fully studied with respect to order, the item-order account is also applicable to other types of organizational principles, as shown by research demonstrating that under appropriate conditions, retrieval is guided by semantic categories rather than order-based information (McDaniel et al., 2000). Language comprehension can be expected to make use of many bases of organization beyond order. Nonetheless, for the case of processing sentences with object RCs, the *order* of the critical NPs seems to provide the most obvious strategy for organizing memory prior to encountering the embedded and main verbs, which provide syntactic and semantic bases for organizing memory.

Below we report two experiments that use eye tracking during reading to study the comprehension of RC sentences in which the frequency of the nouns in the critical NPs have been manipulated in order to determine whether the tradeoffs in encoding of items and order that are captured by the item-order account can be extended to processes observed in online sentence processing. In Experiment 1, we examine the effect of RC type on sentence comprehension when the two critical NPs in a sentence are either both high frequency or both low frequency, a stimulus configuration that makes each sentence analogous to a pure list in the memory literature. In Experiment 2, we examine the effect of RC type on sentence comprehension when the two critical NPs differ in frequency, a stimulus configuration that is analogous to a mixed list in the memory literature.

Experiment 1

In this experiment, participants' eye movements were recorded while they read sentences that contained either a subject RC (SRC) or an object RC (ORC). The critical NPs of the sentence were always definite descriptions, and the NPs within a given sentence were both either high frequency (HF) or low frequency (LF). In combination, the RC-type and word-frequency manipulations yielded four types of experimental sentences, as illustrated in 4A–4D.

4A. SRC($NP1_{HF}$ – $NP2_{HF}$): The politician that bothered the guest caused a big scandal.

4B. ORC(NP1$_{HF}$ – NP2$_{HF}$): The politician that the guest bothered caused a big scandal.

4C. SRC(NP1$_{LF}$ – NP2$_{LF}$): The handyman that bothered the machinist caused a big scandal.

4D. ORC(NP1$_{LF}$ – NP2$_{LF}$): The handyman that the machinist bothered caused a big scandal.

Manipulation of word frequency should provide evidence that lexical encoding of high-frequency words occurs more readily than lexical encoding of low-frequency words. Manipulation of RC type should provide evidence for greater processing difficulty for sentences containing object RCs, as compared with subject RCs, an amply demonstrated difference that has been shown using a variety of measures. Beyond these expected main effects, evaluating the presence and nature of an interaction between RC type and word frequency allows a first step in evaluating whether the item-order framework, developed to account for memory phenomena observed for lists of words, can be fruitfully extended to memory-dependent processes that occur during sentence comprehension.

Memory-based accounts of the object–subject asymmetry attribute the difference in processing to the greater demands that object RCs place on memory retrieval, as compared with subject RCs. In particular, we believe that a straightforward extension of the item-order framework to sentence processing predicts that the object–subject difference should be reduced for sentences in which the two critical nouns are both high frequency, as compared with when both are low frequency. According to this framework, pure lists of high-frequency words promote more successful encoding of order-based information than do pure lists of low-frequency words, and this information is used to boost retrieval performance (DeLosh & McDaniel, 1996). For object RCs in English, the order of the two initial NPs provides the critical information about their syntactic functions with respect to the embedded and main verbs. As such, fast and accurate retrieval of this order information constitutes a potential memory constraint on the process of the relations within the sentence that are specified by the embedded and main verbs.

Eye tracking during reading provides a method for localizing the effects of word frequency and type of RC on reading comprehension. A great deal of research with this method has shown that high-frequency words take less time to encode and identify than do low-frequency words. The effect of word frequency is seen in first-pass reading times on a word (e.g., Inhoff & Rayner, 1986; Rayner & Duffy, 1986; see Rayner, 1998, for a review), as well as in the probability that a word will be skipped during a first-pass reading (e.g., Henderson & Ferreira, 1993; Rayner, Sereno, & Raney, 1996). The effects of word frequency sometimes "spill over" on first-pass reading times for the immediately following word (e.g., Rayner & Duffy, 1986; Rayner, Sereno, Morris, Schmauder, & Clifton, 1989), a finding attributed to how the opportunity to process the next word depends on the difficulty of identifying the current word (Pollatsek, Juhasz, Reichle, Machacek, & Rayner, 2008; Pollatsek, Reichle, & Rayner, 2006; Reichle, Pollatsek, Fisher, & Rayner, 1998). Thus, during reading, the word-frequency effect is a local one that influences reading time during fixations on that word and, occasionally, the next word.

The locus (or loci) of the greater processing difficulty associated with object RCs, as compared with subject RCs, has also been investigated with eye tracking during reading (and other online methods). Doing so presents the challenge that although the same words can be put in sentences with object and subject RCs, those words appear in a different order. One long-standing perspective on this issue is that processing difficulty begins at the embedded verb (the last word of the object RC) and that it continues on the main-clause verb (which is the same for both types of sentences). As noted above, correct understanding of the sentential relations specified by these verbs depends on memory for the two initial NPs. Initial evidence for this localization was obtained by Holmes and O'Regan (1981) in an eye tracking experiment on French RCs and by Ford (1983) using a continuous lexical decision task on the individual words in a sentence. Both of these articles justified the comparison of performance on different words (the verb at the end of an object RC with the noun at the end of the subject RC) by noting that there was a clear Position × Word-Class interaction, such that there was no difference between these same words when they were the first word in the embedded clause (a verb for a subject RC and a noun for an object RC). King and Just (1991) endorsed this perspective in analyzing results from word-by-word, self-paced reading experiments, presenting single-word reading times for the last word of the RC and for the main clause verb. These reading times were longer for object RCs than subject RCs, whereas average reading time per word for the sentence beginning and the sentence ending did not differ as a function of RC type. In subsequent self-paced reading studies, Gordon et al. (2001) also focused analysis of the sentence complexity effect on the last word of the RC and the main verb of the sentence but, in addition, presented reading times for the individual words in the sentence; those show that the first content word of the RC (verb or noun) is read more quickly than the second content word of the RC (noun or verb) but that this difference is much greater for object RCs than for subject RCs. In additional research with self-paced reading methods (Gordon et al., 2004; Wells, Christiansen, Race, Acheson, & MacDonald, 2009) or eye tracking during reading (Demberg & Keller, 2007; Gordon et al., 2006; Traxler et al., 2002, 2005), researchers have examined either reading time on the last word of the RC and on the main verb or over the entire RC and the main verb.

Although the evidence discussed above has characterized the processing difficulty for object RCs as occurring at points in the sentence where the previously seen nouns must be retrieved (the verbs), in two recent articles, researchers have argued that processing difficulty for object RCs begins much earlier at the determiner ("the") for the embedded noun that begins the object RC. Forster, Guerrera, and Elliot (2009) showed this using a "maze task" in which participants had to choose which of two words provided the best continuation to the preceding words on the trial. Staub (2010) assessed the effects of type of RC by comparing eye-movement patterns and times for the same words across the different positions in which they occur in object and subject RCs. The results showed a higher rate of first-pass regressions from the determiner and the noun in an object RC than in a subject RC. Staub (2010) argued that when taken together with the pattern of reading times, this finding showed that in addition to the later point of processing difficulty associated with memory retrieval, there was an earlier point of processing difficulty that could be attributed

to violations of expectations about how a sentence is likely to proceed. Alternatively, we hypothesize that effects of RC type that occur early in the sentence may represent increased effort at memory encoding in anticipation of later need for retrieval.

Method

Participants. Forty students at the University of North Carolina at Chapel Hill served as participants in this experiment. They were all native English speakers and received course credit for their participation. All participants had normal or corrected-to-normal vision.

Materials. During each experimental session, participants were presented with 36 experimental sentences and 44 filler sentences. Twenty-four of the experimental sentences were adapted from Gordon et al. (2006), and an additional 12 sentences were created for this experiment. As illustrated in 4A–4D, all sentences contained a definite description in both the head NP position (NP1) and in the embedded NP position (NP2). Word frequency estimates were derived from the Celex 2 frequency database (Baayen, Piepenbrock, & Gulikers, 1995). The logarithms of the total number of occurrences per million for the high-frequency descriptions were all greater than 1.3, whereas the logarithms for the total number of occurrences per million of the low-frequency descriptions were all less than 0.5. The groups of words were similar in length (mean number of characters for both high- and low-frequency words = 7.5) and number of syllables (low-frequency = 2.6; high-frequency = 2.5). Four versions of each sentence were created by crossing RC type and NP frequency across participants; further, an additional four versions of each sentence were created by swapping the positions of the two definite descriptions so that any differences in semantic congruity between NPs and the different argument roles for the embedded verb were balanced across subject RC and object RC constructions. Accordingly, there were eight different counterbalancing lists, and each participant was presented with one of these lists. Appendix A lists all experimental stimuli in their object-extracted form. The filler sentences contained a variety of syntactic structures but contained no RCs. After each sentence, a comprehension question about the content of the sentence was presented. Following the procedures of King and Just (1991), two thirds of the comprehension questions presented after the experimental sentences asked about the action being described in the RC, and the other third asked about the action being described in the main clause. Similar sorts of comprehension questions appeared after each filler sentence. Half of all the comprehension questions were true, and half were false.

Design and procedure. Within each experimental session, there were four blocks. The first block contained 14 filler sentences and served as a warm-up block. The next three blocks each contained 10 filler sentences and 12 experimental sentences. The order of sentence presentation was randomized within each block. Participants were instructed to read naturally and to press the space bar after reading each sentence, at which time the comprehension question appeared. Participants pressed one key to answer *true*, and another key to answer *false*.

Throughout the experiment, each participant wore an EyeLink I system eye-tracking device that was distributed by SensoMotoric Instruments (SMI). The eye tracker sampled pupil location at a rate of 250 Hz. In addition, the system parsed the samples into fixations and saccades. After undergoing a routine that calibrated the eye tracker, participants began the experimental run. The materials of the experiment were presented on a computer screen with software provided by SMI. Each trial began with the presentation of a fixation point on the screen at the location where the first word of the sentence would later be presented. The presentation of this fixation point served both to direct the gaze of the participant to the location of the beginning of the sentence and to maintain the calibration of the eye tracker. During the presentation of the fixation point, the experimenter used another computer to monitor the location of the participant's gaze. When the participant's gaze was judged to be sufficiently steady on the fixation point, the experimenter pressed a button to make the fixation point disappear and the sentence appear. Sentences were presented in 20-point Times New Roman font, and the critical regions of the sentence always fit on the first line. After reading the sentence, the participant pressed the spacebar. The sentence then disappeared and the comprehension question relating to that sentence appeared. After the participant pressed the button corresponding to his or her answer to the question, the trial ended, and the fixation point for the next trial appeared. During each trial, the experimenter could see the location of the participant's fixation relative to the location of the words of the trial on the computer that the experimenter was using. If the calibration of the eye tracker appeared inadequate, the experimenter recalibrated the eye tracker between trials.

Results

Short fixations (less than 80 ms) were combined with an adjacent fixation if it fell on the same word (Rayner, 1978); otherwise, they were eliminated. Fixations longer than 800 ms were trimmed to 800 ms (Kwon, Lee, Gordon, Kluender, & Polinsky, 2010; Lee, Lee, & Gordon, 2007). These criteria were also applied in Experiment 2.

Analyses of the data relied on two eye tracking measures of reading time, gaze duration, and regression-path duration, as well as the offline measure of comprehension-question accuracy. *Gaze duration* is defined as the sum of the durations of the initial fixations on a word, provided that no material downstream in the sentence has been viewed; it terminates when the gaze is first directed away from the word, regardless of whether the subsequent fixation is progressive or regressive. As is conventional, calculation of mean gaze duration across trials excluded trials in which the word was skipped on first-pass reading. *Regression-path duration* (also called *go-past time*) is the sum of all fixations from the first fixation on a word up to, but excluding, the first fixation downstream from this region. Additional analyses of word targeting are presented in Appendix B.

Reading times. Reading times from all trials were included, regardless of whether the comprehension question was answered correctly. Table 1 shows reading-time results (gaze duration and regression-path duration) for the critical content words in the two types of relative-clause sentences. We focus on the content words because the function words (the complementizer—*that*—and the determiner—*the*) are frequently skipped during reading, which means that reading times often reflect only a very limited portion of the data (see Appendix B). In addition, for the head noun we only report gaze duration because the only word preceding it is the determiner, providing little space for regressions. Analyses of the

Table 1
Eye-Tracking Results of Experiment 1: Reading-Time Measures for Content Words

Frequency pattern	Example word	Gaze	Regression path	Example word	Gaze	Regression path
	Object RC head noun			Subject RC head noun		
LF–LF	*handyman*	330		*handyman*	329	
HF–HF	*politician*	262		*politician*	273	
	Object RC embedded noun			Subject RC embedded verb		
LF–LF	*machinist*	299	479	*bothered*	289	437
HF–HF	*guest*	247	402	*bothered*	260	379
	Object RC embedded verb			Subject RC embedded noun		
LF–LF	*bothered*	366	541	*machinist*	348	513
HF–HF	*bothered*	333	485	*guest*	302	391
	Object RC main verb			Subject RC main verb		
LF–LF	*caused*	343	569	*caused*	339	522
HF–HF	*caused*	332	550	*caused*	323	479

Note. RC = relative clause; LF = low frequency; HF = high frequency.

effects of RC type on the head noun and main-clause verb are relatively straightforward because these words occur in the same positions across the two types of sentences. In contrast, analyses of the effects of RC type on the noun and verb that are embedded in the RC are complicated because the positions of the two words differ for the two types of RCs (i.e., the embedded noun precedes the embedded verb in an object RC, but the embedded verb precedes the embedded noun in an subject RC). Our analyses of the effects of RC type on reading times for the embedded noun and verb examine the main effects of position and RC type and the interaction of these factors.

Head noun (N1). There was a strong effect of word frequency on early measures of processing, as indicated by our measure of gaze duration. High-frequency N1s were read significantly faster (267 ms) than low-frequency N1s (329 ms), $F_1(1, 39) = 21.96$, $MSE = 6,968$, $p < .001$; $F_2(1, 35) = 40.83$, $MSE = 3,527$, $p < .001$. There was no main effect of RC type on gaze durations on N1, $F_1(1, 39) < 1$; $F_2(1, 35) < 1$, and no interaction between noun frequency and RC type, $F_1(1, 39) < 1$; $F_2(1, 35) < 1$. This pattern of nonsignificant results would be expected, given that no text downstream from N1 had been fixated yet. Although participants may have received some parafoveal information about upcoming words in the sentence, we would not expect this information to be strong enough to influence reading time on the head noun.

Embedded noun (N2). As with N1, we observed a strong main effect of N2 frequency on gaze duration, with high-frequency N2s being read significantly faster (274 ms) than low-frequency N2s (323 ms), $F_1(1, 39) = 28.34$, $MSE = 3,386$, $p < .001$; $F_2(1, 35) = 28.57$, $MSE = 3,125$, $p < .001$. The magnitude of the word frequency effect did not depend on whether the embedded noun was the first constituent in the RC (as in an object RC) or the second constituent (as in a subject RC), $F_1(1, 39) < 1$; $F_2(1, 35) < 1$. The measure of regression-path duration also showed this pattern of effects on reading times of the embedded noun, with shorter durations for high-frequency N2s (397 ms) than for low-frequency N2s (496 ms), $F_1(1, 39) = 27.73$, $MSE = 14,203$, $p < .001$; $F_2(1, 35) = 31.56$, $MSE = 11,566$, $p < .001$, but no dependence of this

effect on RC type, $F_1(1, 39) = 1.57$, $MSE = 12,856$, $p > .21$; $F_2(1, 35) = 1.3$, $MSE = 13,894$, $p > .26$.

Content words in the RC. Gaze durations showed no main effect of RC type for the two embedded content words, $F_1(1, 39) = 1.29$, $MSE = 8,153$, $p > .26$; $F_2(1, 35) = 3.15$, $MSE = 3,625$, $p > .08$. In addition, there was no significant interaction of RC type with NP frequency, $F_1(1, 39) < 1$; $F_2(1, 35) < 1$. Gaze duration did show a highly significant effect of word position within the RC, with the last content word of the RC (337 ms) being read more slowly than the first (273 ms), $F_1(1, 39) = 108.96$, $MSE = 2,976$, $p < .001$; $F_2(1, 35) = 67.37$, $MSE = 4359$, $p < .001$. These patterns were examined more closely with a pair of contrasts that holds position within the RC constant (see Appendix B for alternative contrasts). For the initial position in the RC, gaze durations were not noticeably different for object RCs (273 ms for the embedded noun), as compared with subject RCs (274 ms for the embedded verb), $t_1(39) = 0.11$, $p > .91$; $t_2(35) = 0.17$, $p > .86$. For the final position, gaze durations were slightly, but significantly, longer for object RCs (349 ms for the embedded verb), as compared with subject RCs (325 ms for the embedded noun), $t_1(39) = 2.10$, $p < .05$; $t_2(35) = 2.03$, $p = .05$.

Regression-path durations showed a main effect of RC type, such that reading times were slower for the two embedded content words in an object RC than a subject RC, $F_1(1, 39) = 7.69$, $MSE = 22,613$, $p < .01$; $F_2(1, 35) = 12.15$, $MSE = 12,719$, $p < .001$. However, there was no significant interaction of RC type with NP frequency, $F_1(1, 39) < 1$; $F_2(1, 35) < 1$. As with gaze duration, regression-path duration showed a significant main effect of word position within the RC, with the last content word (482 ms) being read more slowly than the first content word (424 ms), $F_1(1, 39) = 10.81$, $MSE = 24,892$, $p < .001$; $F_2(1, 35) = 15.60$, $MSE = 16,535$, $p < .001$. For the initial position in the RC, there was no significant difference in regression-path durations for object RCs (439 ms), as compared with subject RCs (408 ms), $t_1(39) = 1.62$, $p > .11$; $t_2(35) = 1.34$, $p > .18$. For the final position, regression-path durations were longer for object RCs (513 ms) than subject RCs (453 ms), $t_1(39) = 2.47$, $p < .02$; $t_2(35) = 2.46$, $p < .02$.

Main-clause verb. Gaze duration on the main-clause verb showed no effects of RC type, $F_1(1, 39) < 1$; $F_2(1, 35) < 1$, frequency, $F_1(1, 39) = 1.95$, $MSE = 3920$, $p > .17$; $F_2(1, 35) = 1.83$, $MSE = 3893$, $p > .18$, or the interaction of those two factors, $F_1(1, 39) < 1$; $F_2(1, 35) < 1$. However, regression-path duration revealed shorter reading times for main-clause verbs following a subject RC (500 ms) than following an object RC (559 ms), $F_1(1, 39) = 4.47$, $MSE = 31,105$, $p < .05$; $F_2(1, 35) = 5.55$, $MSE = 28,569$, $p < .03$. Regression-path duration at the main-clause verb showed no main effect of NP frequency, $F_1(1, 39) = 1.12$, $MSE = 36,026$, $p > .29$; $F_2(1, 35) = 2.87$, $MSE = 27,684$, $p > .09$, and no interaction between NP frequency and RC type, $F_1(1, 39) < 1$; $F_2(1, 35) < 1$.

Summary and interpretation of reading-time results. The reading-time results provide evidence about three ways in which the experimental manipulations affected reading comprehension:

1. As expected from previous research, high-frequency nouns were encoded more rapidly than low-frequency nouns, as shown by the highly significant effects of word frequency on gaze duration for the manipulated nouns (both in the main and embedded clauses).

2. Again, as expected from previous research, sentences with object RCs were processed more slowly than were sentences with subject RCs. This effect of the type of RC was superimposed on a general slowing of reading for the final word for both types of RCs and extended to the main verb of the sentence, a pattern shown in Holmes and O'Regan (1981); Ford (1983), and Gordon et al. (2001), and discussed by King and Just (1991).

3. The magnitude of the slowdown in processing for object RCs, as compared with subject RCs, did not depend on the frequency of the nouns in the sentences.

Comprehension-question accuracy. Table 2 shows mean comprehension-question accuracies. Because these values were all relatively close to the upper limit of the distribution, the data were arcsine transformed prior to calculation of inferential statistics (Cohen, Cohen, West, & Aiken, 2003). There was a main effect of RC type, with responses to questions following subject RCs being significantly more accurate (91.4%) than questions following object RCs (86.4%), $F_1(1, 39) = 7.72$, $MSE = 0.19$, $p < .01$; $F_2(1, 35) = 10.25$, $MSE = 0.16$, $p < .005$. There was no main effect of NP frequency on comprehension question accuracy, $F_1(1, 39) < 1$; $F_2(1, 35) < 1$, nor was there an interaction between NP frequency and RC type, $F_1(1, 39) = 1.47$, $MSE = 0.29$, $p > .20$; $F_2(1, 35) < 1$.

Table 2
Mean Comprehension Question Accuracies for Experiment 1

| RC type | Frequency | | Mean accuracy % |
	NP1	NP2	
Subject RC	LF	LF	89.4
Object RC	LF	LF	86.1
Subject RC	HF	HF	93.3
Object RC	HF	HF	86.7

Note. RC = relative clause; LF = low frequency; HF = high frequency; NP1 = head noun phrase; NP2 = embedded noun phrase.

Discussion

In line with previous studies, the results of Experiment 1 showed strong effects of word frequency on eye-movement measures of word recognition, such that high-frequency words, regardless of their syntactic context, received shorter fixation times than low-frequency words. Also in line with previous studies, the results showed that object RCs caused greater processing difficulty than subject RCs, as indicated by longer regression-path durations at the RC region of the sentence, as well as the main-clause verb. We also observed this main effect of RC type on the offline measure of comprehension-question accuracy.

The absence of interactions between word frequency and sentence type on measures of reading time and comprehension-question accuracy provides no evidence that word frequency has effects that extend beyond local word processing to the higher level language processes that are involved in understanding sentences with object RCs. Accordingly, the results provide no support for our attempt to extend the item-order account to the processing of complex sentences. Our extension was based on the premise that interpretation of an object RC sentence requires accurate memory for the order of the first two NPs and on the finding in the list-memory literature that order is best recalled in pure lists of high-frequency items. This reasoning leads to the prediction that the object–subject difference should be smaller when both critical NPs are high frequency than when both are low frequency, a pattern that was not observed. Although this experiment yielded a negative result with respect to how the order-focused processes of the item-order account contribute to the processing of complex sentences, it provides a pure-list baseline for determining whether list-composition effects are observed in sentence processing, and it leaves open the possibility that item-focused processes specified by the item-order account have an impact on sentence processing. These issues are pursued in the next experiment.

Experiment 2

Experiment 2 investigated the processing of RC sentences in which the word frequency of the two critical NPs differed (i.e., one was low-frequency and the other was high-frequency), a type of sentence composition analogous to the mixed lists used in memory research. In contrast, Experiment 1 used RC sentences where the critical NPs were of the same type (both high frequency or both low frequency), a situation analogous to pure lists used in memory research. As shown in 5A–5D, Experiment 2 had four conditions generated from the sentences used in Experiment 1.

5A.   SRC(NP1$_{HF}$ – NP2$_{LF}$): The guest that bothered the machinist caused a big scandal.

5B.   ORC(NP1$_{HF}$ – NP2$_{LF}$): The guest that the machinist bothered caused a big scandal.

5C.   SRC(NP1$_{LF}$ – NP2$_{HF}$): The machinist that bothered the guest caused a big scandal.

5D.   ORC(NP1$_{LF}$ – NP2$_{HF}$): The machinist that the guest bothered caused a big scandal.

Examination of whether—and how—these orderings of word frequency affect the magnitude of the object–subject difference in ease of processing provides a second test of whether the frequency of critical NPs influences higher level sentence processing. If the results show such a moderating effect, then the contrast between the effects of word frequency in this experiment, as compared with the previous one, would provide evidence for a sentence-composition effect that it is analogous to the list-composition effects found in the memory literature.

Experiment 1 tested the prediction that the ease of comprehending sentences with object RCs is critically limited by accuracy of memory for the order of the two critical NPs. Research with long-term memory paradigms provides ample support for the idea that order is better recalled for high-frequency items than for low-frequency items, though this effect has not been found in short-term memory paradigms with immediate serial recall (Hulme et al., 2003). As discussed above, the results of Experiment 1 provided no support for this prediction. Here, we consider an alternative possibility: that ease of comprehension is limited by ease of item recall specifically for the head of the RC (NP1). Ease of recalling the embedded noun in an object RC should not influence ease of comprehension because it is immediately available (cost free) in working memory (McElree, 2006; McElree, Foraker, & Dyer, 2003).

The list-composition literature provides consistent evidence that unusual (e.g., infrequent) words are recalled better than common (e.g., frequent) words when the two types of words are presented together in a mixed list (DeLosh & McDaniel, 1996; McDaniel & Bugg, 2008). If sentences with object RCs are harder to comprehend than ones with subject RCs because NP1 must be retrieved on encountering the embedded verb, then extension to sentence processing of the item-processing mechanisms of the item-order account predicts that facilitation of recall for a low-frequency NP1 (paired with a high-frequency NP2) should reduce the object–subject difference, as compared with that found for a high-frequency NP1 (paired with a low-frequency NP2). Thus, this account predicts that the object–subject difference ought to be reduced when the critical NPs have the order LF–HF (e.g., 5C & 5D), as compared with the order HF–LF (e.g., 5A & 5B).

Method

Participants.   Forty students at the University of North Carolina at Chapel Hill served as participants in this experiment. They were all native English speakers and received course credit for their participation. All participants had normal or corrected-to-normal vision.

Materials, design, and procedure.   The 44 filler sentences that were used in Experiment 1 were also used here. The critical NPs of the experimental sentences from Experiment 1 were recombined so that each sentence contained one critical NP that was low frequency and one critical NP that was high frequency. The ordering of these NPs with respect to their frequency was additionally counterbalanced so that each recombination of the experimental sentences was presented with the high-frequency noun serving as NP1 in one list and the low-frequency noun serving as NP1 in another list. Thus, this experiment consisted of eight counterbalanced lists that were determined by the combination of frequency ordering and RC type for a set of four definite descriptions. All other aspects of the design and procedure were identical to that of Experiment 1.

Results

The approach used to analyze the data was the same as that of Experiment 1, focusing in turn on reading time patterns and comprehension-question accuracy (Appendix C provides additional analyses).

Reading times.   All trials were analyzed, regardless of whether the comprehension question was answered correctly. Table 3 shows the reading-time results for the critical content words in the two types of relative-clause sentences.

Head noun (N1).   As in Experiment 1, there was a strong effect of the frequency of N1, with high-frequency nouns being

Table 3
Eye-Tracking Results of Experiment 2: Reading-Time Measures for Content Words

Frequency pattern	Example word	Gaze	Regression path	Example word	Gaze	Regression path
		Object RC head noun			Subject RC head noun	
LF–HF	*machinist*	333		*machinist*	342	
HF–LF	*guest*	292		*guest*	272	
		Object RC embedded noun			Subject RC embedded verb	
LF–HF	*guest*	253	382	*bothered*	289	393
HF–LF	*machinist*	297	467	*bothered*	260	386
		Object RC embedded verb			Subject RC embedded noun	
LF–HF	*bothered*	351	571	*guest*	310	459
HF–LF	*bothered*	369	641	*machinist*	341	471
		Object RC main verb			Subject RC main verb	
LF–HF	*caused*	353	688	*caused*	321	544
HF–LF	*caused*	341	609	*caused*	339	536

Note.   RC = relative clause; LF = low frequency; HF = high frequency.

read significantly more quickly (282 ms) than low-frequency nouns (338 ms), $F_1(1, 39) = 14.52$, $MSE = 8,475$, $p < .001$; $F_2(1, 35) = 20.17$, $MSE = 5,320$, $p < .001$. As would be expected, gaze duration was not affected by RC type, $F_1(1, 39) < 1$; $F_2(1, 35) < 1$, nor was there an interaction between frequency and RC type, $F_1(1, 39) = 1.70$, $MSE = 5,221$, $p > .20$, $F_2(1, 35) = 1.40$, $MSE = 6,511$, $p > .24$.

Embedded noun (N2). As with N1, gaze duration on N2 showed a strong effect of word frequency, with high-frequency N2s being read significantly faster (281 ms) than low-frequency N2s (319 ms), $F_1(1, 39) = 16.87$, $MSE = 3,315$, $p < .001$; $F_2(1, 35) = 10.34$, $MSE = 5,006$, $p < .004$. The magnitude of the word frequency effect did not depend on whether the embedded noun was the first constituent in the RC (as in an object RC) or the second constituent (as in a subject RC), $F_1(1, 39) < 1$; $F_2(1, 35) < 1$. Analyses of regression-path duration showed the same pattern, a main effect of frequency (marginal in the item analysis), $F_1(1, 39) = 5.74$, $MSE = 16,407$, $p < .03$; $F_2(1, 35) = 3.42$, $MSE = 25,791$, $p < .07$, and no interaction between frequency and position of the embedded noun, $F_1(1, 39) = 2.45$, $MSE = 21,436$, $p > .12$, $F_2(1, 35) = 1.88$, $MSE = 24,539$, $p > .17$.

Content words in the RC. Gaze durations on the two embedded content words were marginally shorter in object RCs, as compared with subject RCs, $F_1(1, 39) = 2.85$, $MSE = 4,720$, $p < .10$; $F_2(1, 35) = 5.89$, $MSE = 2,112$, $p < .03$. Gaze durations were longer for the final than the initial content word of the RC, $F_1(1, 39) = 61.39$, $MSE = 5,261$, $p < .001$; $F_2(1, 35) = 125.81$, $MSE = 2,276$, $p < .001$. Contrasts controlling for position within the RC showed that gaze durations did not differ in the initial position (275 ms for the noun embedded in an object RC vs. 283 ms for the verb embedded in a subject RC), $t_1(39) = 0.81$, $p > .42$; $t_2(35) = 0.63$, $p > .53$. For the final position, gaze durations were longer for the verb in an object RC (360 ms), as compared with the noun in a subject RC (325 ms), $t_1(39) = 2.73$, $p < .01$; $t_2(35) = 2.90$, $p < .01$ (see Appendix C for alternative contrasts).

Regression-path durations for the embedded content words were longer in object RCs than subject RCs, $F_1(1, 39) = 33.11$, $MSE = 18,769$, $p < .001$; $F_2(1, 35) = 29.75$, $MSE = 18,842$, $p < .001$. Critically, this effect of RC type interacted significantly with the frequency manipulation such that the object–subject asymmetry was smaller when N1 was low-frequency and N2 was high-frequency than when N1 was high-frequency and N2 was low-frequency, $F_1(1, 39) = 8.55$, $MSE = 13,148$, $p < .007$; $F_2(1, 35) = 5.48$, $MSE = 16,943$, $p < .03$. Finally, there was a significant effect of position, with longer reading times on RC-final content words than RC-initial content words, $F_1(1, 39) = 21.01$, $MSE = 62,867$, $p < .001$; $F_2(1, 35) = 58.22$, $MSE = 19,880$, $p < .001$. The contrasts controlling for word position showed that in initial position, regression-path durations trended toward being slower for nouns in object RCs (426 ms) than verbs in subject RCs (390 ms), $t_1(39) = 2.17$, $p < .04$; $t_2(35) = 1.85$, $p < .08$, and that in final position, times were significantly longer for the verb embedded in object RCs (607 ms), as compared with the noun embedded in subject RCs (465 ms), $t_1(39) = 5.99$, $p < .001$; $t_2(35) = 3.96$, $p < .001$.

Main-clause verb. Analysis of gaze duration on the main-clause verb revealed a marginally significant main effect of RC type, with shorter times for verbs following subject RCs (330 ms) than verbs following object RCs (347 ms), $F_1(1, 39) = 3.10$, $MSE = 3,949$, $p < .09$; $F_2(1, 35) = 3.78$, $MSE = 2,763$, $p < .07$. There was no main effect of the frequency manipulation, $F_1(1, 39) = 0.08$, $MSE = 4,646$, $p > .78$; $F_2(1, 35) = 0.01$, $MSE = 5,367$, $p > .91$; the interaction between NP frequency and RC type was not significant in the subject analysis but was in the item analysis, $F_1(1, 39) = 2.63$, $MSE = 3,464$, $p > .11$, $F_2(1, 35) = 4.2$, $MSE = 2,345$, $p < .05$. Regression-path duration showed shorter durations for verbs following subject RCs (540 ms) than for verbs following object RCs (649 ms), $F_1(1, 39) = 8.92$, $MSE = 52,818$, $p < .006$; $F_2(1, 35) = 8.69$, $MSE = 69,190$, $p < .007$. There was no main effect of NP frequency, $F_1(1, 39) = 1.34$, $MSE = 57,232$, $p > .25$; $F_2(1, 35) = 1.19$, $MSE = 72,540$, $p > .28$, nor was there a significant interaction between NP frequency and RC type, $F_1(1, 39) = 0.75$, $MSE = 66,565$, $p > .39$; $F_2(1, 35) = 0.65$, $MSE = 75,759$, $p > .42$.

Summary and interpretation of reading-time results. As with the previous experiment, the reading-time patterns provide evidence about three ways in which the experimental manipulations affected reading comprehension: 1. High-frequency nouns were encoded more rapidly than low-frequency nouns in both the main-clause and embedded-clause roles. 2. Sentences with object RCs were processed more slowly than were ones with subject RCs, an effect that began at the final word of the embedded clause and continued through the main verb of the sentence. 3. The magnitude of the slowdown in processing for object RCs, as compared with subject RCs, was reduced for regression-path duration when the head noun was low frequency and the embedded noun was high frequency, as compared with the reverse configuration. Although the first two patterns (frequency and the general effect of relative-clause processing) essentially repeat findings from Experiment 1, the third finding is new and is critical to the goal of determining whether word frequency contributes to higher level aspects of sentence processing and whether its contribution is moderated by a sentence-composition effect analogous to the list-composition effects found in the memory literature. The interaction of RC type and frequency sequence was found on regression-path duration measures for the content words in the RC.

As no three-way interactions were found between RC type, frequency sequence, and position within the RC, it is not possible to localize this effect precisely as occurring on the initial or final content word of the RC. Because retrieval processes for interpreting the object RC are unlikely to occur before the final content word of the RC, an additional analysis was performed on total reading time on the words in the RC, given that the final word of the RC had been fixated. Times were faster for object RCs in the LF–HF condition (1,382 ms) than in the HF–LF condition (1,533), whereas for subject RCs, times were somewhat slower in the LF–HF condition (1,315) than in the HF–LF condition (1,185), $F_1(1, 39) = 9.05$, $MSE = 87,073$, $p < .006$; $F_2(1, 35) = 10.15$, $MSE = 70,759$, $p < .004$. Contrasts showed that the effect was significant for object RCs, $t_1(39) = 2.28$, $p < .05$; $t_2(35) = 2.30$, $p < .05$, but not for subject RCs, $t_1(39) = 1.88$, $p > .05$; $t_2(35) = 1.68$, $p > .05$.

Comprehension-question accuracy. Table 4 shows mean accuracy on the comprehension questions broken down by condition. As in Experiment 1, data were arcsine transformed before inferential statistics were calculated. There was a trend toward a main effect of RC type, with comprehension question accuracy higher in response to questions following sentences with subject

Table 4

Mean Comprehension Question Accuracies for Experiment 2

RC type	Frequency		Mean accuracy %
	NP1	NP2	
Subject RC	LF	HF	92.2
Object RC	LF	HF	91.9
Subject RC	HF	LF	91.7
Object RC	HF	LF	82.4

Note. RC = relative clause; LF = low frequency; HF = high frequency; NP1 = head noun phrase; NP2 = embedded noun phrase.

RCs (91.9%) than in response to questions following sentences with object RCs (88.1%). This effect was significant by items, but fell short of significance in our subjects analysis, $F_1(1, 39) = 2.84$, $MSE = 0.27$, $p < .10$; $F_2(1, 35) = 10.27$, $MSE = 0.11$, $p < .004$. Comprehension question accuracies were higher for questions following sentences with the sequence LF–HF (92.1%) than for questions following sentences with the sequence HF–LF (87.9%), $F_1(1, 39) = 7.70$, $MSE = 0.12$, $p < .009$; $F_2(1, 35) = 5.45$, $MSE = 0.15$, $p < .03$. Moreover, the interaction between NP sequence and RC type was significant, $F_1(1, 39) = 8.06$, $MSE = 0.12$, $p < .008$; $F_2(1, 35) = 6.29$, $MSE = 0.12$, $p < .02$. For questions following sentences with object RCs, accuracies were 7.2% higher when the sequence was LF–HF, as compared with when it was HF–LF, $t_1(39) = 3.80$, $p < .001$; $t_2(35) = 3.79$, $p < .001$.

Discussion

Experiment 2 demonstrated that the processing of object RCs, compared with their subject RC counterparts, is facilitated when the head noun is low frequency and the embedded noun is high frequency (LF–HF), as compared with when this frequency configuration is reversed (HF–LF). This critical interaction was found on reading-time measures and also on accuracy in answering comprehension questions. This pattern shows that word frequency not only affects local lexical processing but also affects higher level processes involved in interpreting complex sentences containing object RCs. The contrast between the interaction of word frequency and RC type found in this experiment and the absence of such an interaction in Experiment 1 demonstrate a sentence-composition effect analogous to the list-composition effect that has been found to moderate the influence of word frequency on recall (DeLosh & McDaniel, 1996; McDaniel & Bugg, 2008).

The manner in which the object–subject asymmetry depended on the ordering of the infrequent and frequent nouns is consistent with our extensions of the encoding and retrieval mechanisms specified in the item-order model developed in the memory literature. The item-based processes in the item-order account would facilitate retrieval of the head noun (N1) in an LF–HF sequence because recall of unusual words is facilitated in mixed lists. Although the pattern of facilitation is consistent with this account, its timing suggests that some portion of the facilitation may have occurred before readers encountered the embedded verb, which on most accounts would provide the cue that retrieval was necessary. One approach might explain this early effect as resulting from

violations of experience-based expectations (as in Gennari & Mac-Donald, 2008; Staub, 2010), though to our knowledge, there is no evidence that sentences with object RCs occur more commonly with the order low-frequency common noun followed by high-frequency common noun, as compared with the reverse order. Alternatively, the early effect could reflect encoding differences of the sort that are expected on the item-order account and that have also been considered in accounts of memory interference in RC processing (Gordon et al., 2006). Experiments focused on memory typically control the time available for encoding, and differences in how that time is used are inferred from memory performance. In contrast, readers in this experiment have control over encoding time (as revealed by the time spent fixating words and their rate of progress through a text), so that the elevated times for object RCs with the frequency pattern HF–LF may reflect encoding strategies that derive from the difficulty in remembering such sequences.

General Discussion

In the experiments reported in this article, we used eye tracking during reading to investigate how the processing of complex sentences is influenced by the particular configuration of "common" versus "unusual" critical NPs; commonness was manipulated by varying the lexical frequency of the NPs' head nouns. Experiment 2 showed that the object–subject asymmetry in ease of processing was reduced when the first NP was unusual (low frequency) relative to the second NP. However, in Experiment 1, where the two critical NPs were the same with respect to commonness (both high frequency or both low frequency), the results showed strong effects of RC type on several measures of processing and strong effects of word frequency on local lexical processing but no interaction between these two factors. These results have implications for theoretical efforts to understand the role of memory during language comprehension and show that ideas from the memory literature about how encoding and retrieval strategies are implicitly influenced by stimulus characteristics are applicable to online sentence processing.

Method and Models of Memory—Applications to Sentence Processing

The experiments reported in this article were motivated by the view that memory encoding and retrieval constitute key limiting factors in the comprehension of complex sentences. This general view is shared by a number of current theories of sentence processing (Gordon et al., 2001, 2002; Grodner & Gibson, 2005; Lewis & Vasishth, 2005; Lewis et al., 2006; Van Dyke & Lewis, 2003; Van Dyke & McElree, 2006). The specific manipulations and predictions used in the experiment were based on research showing how the commonness of items (particularly word frequency) and the composition of a list of to-be-remembered items affect performance in memory tasks. Varying the composition of a list of to-be-remembered items has long been a valuable manipulation in research on human memory (e.g., DeLosh & McDaniel, 1996; Duncan, 1974; Gregg, 1976; Gregg et al., 1980; May & Tryk, 1970; Merritt et al., 2006).

In the present research, we applied this manipulation to the study of sentence processing, with the contrast between the results of Experiments 1 and 2 demonstrating a sentence-composition

effect in which the mix of types of NPs in a sentence affected the ease of processing complex sentences as indexed by the object–subject asymmetry. In particular, these experiments showed that the impact on sentence comprehension of the commonness of items is greater in mixed lists consisting of both common and unusual items than it is on pure lists consisting of either common items or unusual items. The item-order account (DeLosh & McDaniel, 1996; McDaniel & Bugg, 2008) explains list composition effects as emerging from the tendency of unusual items such as low-frequency words to focus attention on the encoding of item-specific information at the expense of encoding order-based information. Both the item and order facets of the item-order account can be seen as possible contributors to the processing difficulty found in object RCs, as compared with subject RCs; evaluating those contributions forces a more in-depth examination of how memory for earlier parts of a sentence is used during the process of interpretation.

Object RCs are an unusual type of English sentence in the sense that two NPs must be stored in memory before they can be integrated into the developing sentence interpretation through explicit relational information (in this case specified by the verbs). As the order of those two NPs determines their roles with respect to the verbs, encoding and accurately retrieving the order of the NPs is a plausible candidate strategy for interpreting sentences with object RCs. The item-order account predicts that use of order information is facilitated when items are easily encoded (such as the pairs of high-frequency NPs that appeared together in sentences in Experiment 1). Contrary to this possibility, the magnitude of object–subject processing difference was not reduced for sentences containing high-frequency nouns, as compared with those containing low-frequency nouns. Although this absence of an interaction could be explained by the view that list-composition effects do not influence sentence processing, such a view would be inconsistent with the list effects observed in Experiment 2. Instead, the results of Experiment 1 suggest that word order *per se* is not critical to the developing interpretation of a sentence, even in the case of object RCs. On this view, direct representation of order would play a negligible role in the processing of higher levels of sentence organization, which instead would be based on the many other types of cues that sentences provide for encoding words into organized memories.

Although the results of these experiments are not readily accounted for by a view of sentence comprehension in which memory for the order of the two NPs is critical, they are consistent with an account in which memory for the first NP is critical to comprehension of the object RCs. In particular, the manner in which sentence composition affected the magnitude of the object–subject difference in processing was consistent with patterns found in the memory literature: The asymmetry was reduced when the head noun was unusual (low frequency) and when the embedded noun was common (high frequency). The results of the experiments are in some ways counterintuitive; it might have been expected that processing object RCs would be facilitated when the head noun was high frequency because high lexical frequency typically leads to easier processing. Their consistency with memory findings shows the value of considering findings from the memory literature in studies of sentence processing.

Conclusion

Language processing and memory are interdependent, with language processing usually facilitating memory but with memory occasionally constraining the ease of language processing. The facilitative effect of language on memory is apparent in demonstrations that sentences with 16 or more words can be repeated verbatim with no errors, whereas accurate performance with word lists is limited to five or six words (see Baddeley, 2000, for discussion). The constraining effects of memory on language processing are less apparent, becoming noticeable only in sentences with structures—such as object RCs—in which words must be held in memory before they can be integrated with the rest of the sentence. In the experiments reported here, we examined whether the item-order framework could account for variation in the ease with which such memory-dependent structures are understood. This framework has been used successfully to explain memory performance as reflecting a tradeoff in which having common (nondistinctive) items that are easy to encode allows resources to be focused on encoding the order of items in a list, whereas having noncommon (or distinctive) items, or ones that are difficult to encode, prompts allocation of resources to the encoding of individual items. The results provided no evidence that facilitating the encoding of order through the use of easy-to-encode common items reduced the difficulty of understanding object RCs. From this, we conclude that directly encoding the order of the nouns in an object RC does not play a central role in the correct interpretation of the sentence. In contrast, the results provided evidence that factors causing more elaborate encoding of the head noun by making it less common (or more distinctive) do reduce the difficulty in understanding object RCs. We attribute this reduction in processing difficulty to greater ease in retrieving a noncommon head noun because it has been encoded more completely and, possibly, to direct processes of encoding when readers encounter a difficult-to-remember sequence. This characterization shows how general principles of memory that predict the degree to which items are encoded may operate during language comprehension.

References

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database* (Release 2) [CD-ROM]. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

Baddeley, A. D. (1986). *Working memory.* New York, NY: Oxford University Press.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4,* 417–423. doi:10.1016/S1364-6613(00)01538-2

Bever, T. G. (1974). The ascent of the specious, or there's a lot we don't know about mirrors. In D. Cohen (Ed.), *Explaining linguistic phenomena* (pp. 173–200). Washington, DC: Hemisphere.

Brysbaert, M., Drieghe, D., & Vitu, F. (2005). Word skipping: Implications for theories of eye movement control in reading. In G. Underwood (Ed.), *Cognitive processes in eye guidance* (pp. 53–77). Oxford. England: Oxford University Press.

Caplan, D., Alpert, N., & Waters, G. (1998). Effects of syntactic structure and propositional number on patterns of regional cerebral blood flow. *Journal of Cognitive Neuroscience, 10,* 541–552. doi:10.1162/089892998562843

Caramazza, A., & Zurif, E. (1976). Dissociation of algorithmic and heuristic processes in sentence comprehension: Evidence from aphasia. *Brain and Language, 3,* 572–582. doi:10.1016/0093-934X(76)90048-1

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Erlbaum.

DeLosh, E. L., & McDaniel, M. A. (1996). The role of order information in free recall: Application to the word-frequency effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1136–1146. doi:10.1037/0278-7393.22.5.1136

Demberg, V., & Keller, F. (2007). Eye-tracking evidence for integration cost effects in corpus data. In D. McNamara & G. Trafton (Eds.), *Proceedings of the 29th Annual Conference of the Cognitive Science Society* (pp. 947–952). Nashville, TN: Cognitive Science Society.

Drieghe, D., Rayner, K., & Pollatsek, A. (2005). Eye movements and word skipping during reading revisited. *Journal of Experimental Psychology: Human Perception and Performance, 31,* 954–969. doi:10.1037/0096-1523.31.5.954

Duncan, C. P. (1974). Retrieval of low-frequency words from mixed lists. *Bulletin of the Psychonomic Society, 4,* 137–138.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102,* 211–245. doi:10.1037/0033-295X.102.2.211

Ford, M. (1983). A method for obtaining measures of local parsing complexity throughout sentences. *Journal of Verbal Learning and Verbal Behavior, 22,* 203–218. doi:10.1016/S0022-5371(83)90156-1

Forster, K. I., Guerrera, C., & Elliot, L. (2009). The maze task: Measuring forced incremental sentence processing time. *Behavior Research Methods, 41,* 163–171. doi:10.3758/BRM.41.1.163

Fox, B., & Thompson, S. A. (1990). A discourse explanation of the grammar of relative clauses in English conversation. *Language, 66,* 297–316. doi:10.2307/414888

Gennari, S. P., & MacDonald, M. C. (2008). Semantic indeterminacy in object relative clauses. *Journal of Memory and Language, 58,* 161–187. doi:10.1016/j.jml.2007.07.004

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition, 68,* 1–76. doi:10.1016/S0010-0277(98)00034-1

Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In Y. Miyashita, A. Marantz, & W. O'Neil (Eds.), *Image, language, brain* (pp. 95–126). Cambridge, MA: MIT Press.

Gordon, P. C., & Hendrick, R. (2005). Relativization, ergativity, and corpus frequency. *Linguistic Inquiry, 36,* 456–463. doi:10.1162/0024389054396953

Gordon, P. C., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 1411–1423. doi:10.1037/0278-7393.27.6.1411

Gordon, P. C., Hendrick, R., & Johnson, M. (2004). Effects of noun phrase type on sentence complexity. *Journal of Memory and Language, 51,* 97–114. doi:10.1016/j.jml.2004.02.003

Gordon, P. C., Hendrick, R., Johnson, M., & Lee, Y. (2006). Similarity-based interference during language comprehension: Evidence from eye tracking during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 1304–1321. doi:10.1037/0278-7393.32.6.1304

Gordon, P. C., Hendrick, R., & Levine, W. H. (2002). Memory-load interference in syntactic processing. *Psychological Science, 13,* 425–430. doi:10.1111/1467-9280.00475

Gregg, V. H. (1976). Word frequency, recognition, and recall. In J. Brown (Ed.), *Recall and recognition* (pp. 183–215). London, England: Wiley.

Gregg, V. H., Montgomery, D. C., & Castano, D. (1980). Recall of common and uncommon words from pure and mixed lists. *Journal of Verbal Learning and Verbal Behavior, 19,* 240–245. doi:10.1016/S0022-5371(80)90202-9

Grodner, D., & Gibson, E. (2005). Consequences of the serial nature of linguistic input. *Cognitive Science, 29,* 261–290. doi:10.1207/s15516709cog0000_7

Gundel, J., Hedberg, H., & Zacarski, R. (1993). Referring expressions in discourse. *Language, 69,* 274–307. doi:10.2307/416535

Hale, J. (2001). A probabilistic Earley parser as a psycholinguistic model. In *Proceedings of North American Chapter of the Association for Computational Linguistics* (Vol. 2, pp. 159–166). Pittsburgh, PA: Association for Computational Linguistics.

Henderson, J. M., & Ferreira, F. (1993). Eye movement control during reading: Fixation measures reflect foveal but not parafoveal processing difficulty. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie, 47,* 201–221. doi:10.1037/h0078814

Holmes, V. M., & O'Regan, J. K. (1981). Eye fixation patterns during the reading of relative-clause sentences. *Journal of Verbal Learning and Verbal Behavior, 20,* 417–430. doi:10.1016/S0022-5371(81)90533-8

Hulme, C., Stuart, G., Brown, G. D. A., & Morin, C. (2003). High- and low-frequency words are recalled equally well in alternating lists: Evidence for associative effects in serial recall. *Journal of Memory and Language, 49,* 500–518. doi:10.1016/S0749-596X(03)00096-2

Hunt, R. R., & Elliott, J. M. (1980). The role of nonsemantic information in memory: Orthographic distinctiveness effects on retention. *Journal of Experimental Psychology: General, 109,* 49–74. doi:10.1037/0096-3445.109.1.49

Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics, 40,* 431–439. doi:10.3758/BF03208203

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99,* 122–149. doi:10.1037/0033-295X.99.1.122

Just, M. A., Carpenter, P. A., Keller, T. A., Eddy, W. F., & Thulborn, K. R. (1996, October). Brain activation modulated by sentence comprehension. *Science, 274,* 114–116. doi:10.1126/science.274.5284.114

King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language, 30,* 580–602. doi:10.1016/0749-596X(91)90027-H

Kwon, N., Lee, Y., Gordon, P. C., Kluender, R., & Polinsky, M. (2010). Cognitive and linguistic factors affecting subject/object asymmetry: An eye-tracking study of pronominal relative clauses in Korean. *Language, 86,* 546–582. doi:10.1353/lan.2010.0006

Ledoux, K., & Gordon, P. C. (2006). Interruption-similarity effects during discourse processing. *Memory, 14,* 789–803. doi:10.1080/09658210600679915

Lee, Y., Lee, H., & Gordon, P. C. (2007). Linguistic complexity and information structure in Korean: Evidence from eye-tracking during reading. *Cognition, 104,* 495–534. doi:10.1016/j.cognition.2006.07.013

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition, 106,* 1126–1177. doi:10.1016/j.cognition.2007.05.006

Lewis, R. L. (1999). Specifying architectures for language processing: Process, control, and memory in parsing and interpretation. In M. Crocker, M. Pickering, & C. Clifton (Eds.), *Architectures and mechanisms for language processing* (pp. 56–89). Cambridge, England: Cambridge University Press. doi:10.1017/CBO9780511527210.004

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science, 29,* 375–419. doi:10.1207/s15516709cog0000_25

Lewis, R. L., Vasishth, S., & Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences, 10,* 447–454. doi:10.1016/j.tics.2006.08.007

Mak, W. M., Vonk, W., & Schriefers, H. (2002). The influence of animacy on relative clause processing. *Journal of Memory and Language, 47,* 50–68. doi:10.1006/jmla.2001.2837

Mak, W. M., Vonk, W., & Schriefers, H. (2006). Animacy in processing relative clauses: The hikers that rocks crush. *Journal of Memory and Language, 54,* 466–490. doi:10.1016/j.jml.2006.01.001

May, R. B., & Tryk, H. E. (1970). Word sequence, word frequency, and

free recall. *Canadian Journal of Psychology/Revue canadienne de psychologie, 24,* 299–304. doi:10.1037/h0082866

McConkie, G. W., Kerr, P. W., Reddix, M. D., & Zola, D. (1988). Eye movement control during reading: I. The location of initial eye fixations on words. *Vision Research, 28,* 1107–1118. doi:10.1016/0042-6989(88)90137-X

McDaniel, M. A., & Bugg, J. M. (2008). Instability in memory phenomena: A common puzzle and a unifying explanation. *Psychonomic Bulletin & Review, 15,* 237–255. doi:10.3758/PBR.15.2.237

McDaniel, M. A., DeLosh, E. L., & Merritt, P. S. (2000). Order information and retrieval distinctiveness: Recall of common versus bizarre material. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1045–1056. doi:10.1037/0278-7393.26.4.1045

McDaniel, M. A., Einstein, G. O., DeLosh, E. L., May, C. P., & Brady, P. (1995). The bizarreness effect: It's not surprising, it's complex. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 422–435. doi:10.1037/0278-7393.21.2.422

McElree, B. (2006). Accessing recent events. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 46, pp. 155–200). San Diego, CA: Academic Press.

McElree, B., Foraker, S., & Dyer, L. (2003). Memory structures that subserve sentence comprehension. *Journal of Memory and Language, 48,* 67–91. doi:10.1016/S0749-596X(02)00515-6

Merritt, P. S., DeLosh, E. L., & McDaniel, M. A. (2006). Effects of word frequency on individual-item and serial order retention: Tests of the order-encoding view. *Memory & Cognition, 34,* 1615–1627. doi:10.3758/BF03195924

Miller, G. A., & Chomsky, N. (1963). Finitary models of language users. In D. R. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. II, pp. 269–321). New York, NY: Wiley.

Mulligan, N. W. (1999). The effects of perceptual interference at encoding on organization and order: Investigating the roles of item-specific and relational information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 54–69. doi:10.1037/0278-7393.25.1.54

Mulligan, N. W. (2002). The generation effect: Dissociating enhanced item memory and disrupted order memory. *Memory & Cognition, 30,* 850–861. doi:10.3758/BF03195771

Nairne, J. S., Riegler, G. L., & Serra, M. (1991). Dissociative effects of generation on item and order retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 702–709. doi:10.1037/0278-7393.17.4.702

Pollatsek, A., Juhasz, B. J., Reichle, E. D., Machacek, D., & Rayner, K. (2008). Immediate and delayed effects of word frequency and word length on eye movements in reading: A reversed delayed effect of word length. *Journal of Experimental Psychology: Human Perception and Performance, 34,* 726–750. doi:10.1037/0096-1523.34.3.726

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the E-Z Reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology, 52,* 1–56. doi:10.1016/j.cogpsych.2005.06.001

Rayner, K. (1978). Eye movements in reading and information processing. *Psychological Bulletin, 85,* 618–660. doi:10.1037/0033-2909.85.3.618

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124,* 372–422. doi:10.1037/0033-2909.124.3.372

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. doi:10.3758/BF03197692

Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R., & Clifton, C. (1989). Eye movements and on-line language comprehension processes. *Language and Cognitive Processes, 4,* Sl21– Sl49. doi:10.1080/01690968908406362

Rayner, K., Sereno, S. C., & Raney, G. E. (1996). Eye movement control in reading: A comparison of two types of models. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 1188–1200. doi:10.1037/0096-1523.22.5.1188

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105,* 125–157. doi:10.1037/0033-295X.105.1.125

Reichle, E. D., Warren, T., & McConnell, K. (2009). Using E-Z Reader to model the effects of higher-level language processing on eye movements during reading. *Psychonomic Bulletin & Review, 16,* 1–21. doi:10.3758/PBR.16.1.1

Schmidt, S. R. (1994). Effects of humor on sentence memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 953–967. doi:10.1037/0278-7393.20.4.953

Serra, M., & Nairne, J. S. (1993). Design controversies and the generation effect: Support for an item-order hypothesis. *Memory & Cognition, 21,* 34–40. doi:10.3758/BF03211162

Staub, A. (2010). Eye movements and processing difficulty in object relative clauses. *Cognition, 116,* 71–86. doi:10.1016/j.cognition.2010.04.002

Traxler, M. J., Morris, R. K., & Seely, R. E. (2002). Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory and Language, 47,* 69–90. doi:10.1006/jmla.2001.2836

Traxler, M. J., Williams, R. S., Blozis, S. A., & Morris, R. K. (2005). Working memory, animacy, and verb class in the processing of relative clauses. *Journal of Memory and Language, 53,* 204–224. doi:10.1016/j.jml.2005.02.010

Van Dyke, J. A., & Lewis, R. L. (2003). Distinguishing effects of structure and decay on attachment and repair: A cue-based parsing account of recovery from misanalyzed ambiguities. *Journal of Memory and Language, 49,* 285–316. doi:10.1016/S0749-596X(03)00081-0

Van Dyke, J. A., & McElree, B. (2006). Retrieval interference in sentence comprehension. *Journal of Memory and Language, 55,* 157–166. doi:10.1016/j.jml.2006.03.007

Vitu, F., O'Regan, J. K., Inhoff, A. W., & Topolski, R. (1995). Mindless reading: Eye-movement characteristics are similar in scanning letter strings and reading texts. *Perception & Psychophysics, 57,* 352–364. doi:10.3758/BF03213060

Wanner, E., & Maratsos, M. (1978). An ATN approach to comprehension. In M. Halle, J. Bresnan, & G. Miller (Eds.), *Linguistic theory and psychological reality* (pp. 119–161). Cambridge England: MIT Press.

Warren, T., & Gibson, E. (2002). The influence of referential processing on sentence complexity. *Cognition, 85,* 79–112. doi:10.1016/S0010-0277(02)00087-2

Warren, T., & McConnell, K. (2007). Investigating effects of selectional restriction violations and plausibility violation severity on eye-movements in reading. *Psychonomic Bulletin & Review, 14,* 770–775. doi:10.3758/BF03196835

Waters, G. S., & Caplan, D. (1996). The capacity theory of sentence comprehension: Critique of Just and Carpenter (1992). *Psychological Review, 103,* 761–772. doi:10.1037/0033-295X.103.4.761

Wells, J. B., Christiansen, M. H., Race, D. S., Acheson, D. J., & MacDonald, M. C. (2009). Experience and sentence processing: Statistical learning and relative clause comprehension. *Cognitive Psychology, 58,* 250–271. doi:10.1016/j.cogpsych.2008.08.002

(Appendices follow)

Appendix A

Stimuli From Experiments 1 and 2

The stimuli from Experiments 1 and 2 are shown below in their object-extracted forms. At the head NP and embedded NP positions, the high-frequency description is listed first, followed by the low-frequency description. The particular combinations of high- and low-frequency words were manipulated as described in the text.

1. The doctor/beautician that the student/hustler praised climbed the mountain just outside of town before it snowed.

2. The teacher/jurist that the officer/bookworm phoned cooked the pork chops in barbecue sauce on New Year's Eve.

3. The leader/shoplifter that the husband/overseer liked dominated the conversation while the game was on television.

4. The minister/peddler that the brother/mailman despised drove the sports car home from work that evening.

5. The daughter/hijacker that the president/jester disliked clipped the coupons out with the dull scissors.

6. The official/swindler that the professor/jogger ignored watched the special about Colombian drugs on the nightly news.

7. The artist/lifeguard that the servant/simpleton insulted read the newspaper article about the fire.

8. The soldier/navigator that the farmer/bellboy admired answered the telephone in the fancy restaurant.

9. The writer/cobbler that the colonel/dietician thanked consulted for many hit movies before 1990.

10. The secretary/ferryman that the director/nominee inspired wrote an autobiography after their friendship became well known.

11. The captain/organist that the reader/mugger distrusted called for help after the restaurant closed.

12. The scientist/gladiator that the owner/anchorman amused made paper dolls out of the newspaper.

13. The chief/gymnast that the player/psychic complimented declined a television interview respectfully.

14. The patient/finalist that the general/moderator questioned wrote a long science fiction novel during the summer vacation.

15. The lawyer/baritone that the visitor/bicyclist recommended changed jobs after the announcement a new merger.

16. The politician/machinist that the guest/handyman described worked in a small building near the bus station.

17. The chairman/looter that the employer/samurai advised recalled the event before the trip got underway.

18. The driver/imposter that the victim/trumpeter criticized talked publicly about the incident after the game.

19. The colleague/innkeeper that the judge/fiddler interviewed had a very small office.

20. The manager/lobbyist that the gentleman/kidnapper called drove a gray truck.

21. The neighbor/gunsmith that the prisoner/abbot contacted spoke very quickly.

22. The partner/suitor that the author/cripple entertained behaved with dignity.

23. The priest/censor that the expert/bowler helped worked in a large foreign bank.

24. The citizen/playmate that the stranger/choirboy envied descended the staircase.

25. The assistant/mariner that the engineer/hooker recruited dragged a heavy suitcase through the crowded airport.

26. The consumer/alchemist that the peasant/curate serenaded visited several family members last Tuesday.

27. The supporter/rapist that the passenger/teller taught constructed a tower of playing cards.

28. The employee/fugitive that the specialist/wrestler tutored carved the turkey at Thanksgiving dinner.

29. The socialist/skeptic that the reporter/botanist adored purchased a pair of shoes.

30. The governor/skipper that the speaker/heretic idolized eavesdropped through the open door.

31. The producer/envoy that the commander/custodian evaluated shivered in the cool wind.

32. The criminal/mentor that the painter/barbarian tolerated poured syrup on the French toast.

33. The sergeant/surveyor that the traveler/cadet pitied coached Little League baseball.

34. The waiter/miser that the economist/rescuer scolded blinked due to a sudden gust of dusty wind.

35. The detective/cashier that the foreigner/forger flattered appreciated the exhibit at the museum.

36. The designer/defector that the guardian/shopper frightened chuckled about the scare in retrospect.

Appendix B

Reading-Time Comparisons of Word Types in Different Positions in the RC for Experiment 1

An alternative approach to handling the covariation between position and word class across object and subject RCs is to hold word class constant and compare reading times across positions (Staub, 2010). These contrasts showed that gaze duration for the embedded noun was shorter when it was in an object RC (273 ms) than when it was in a subject RC (325 ms), $t_1(39) = 4.60$, $p < .001$; $t_2(35) = 8.89$, $p = .001$, but that gaze duration for the embedded verb was longer when it was in an object RC (349 ms) than when it was in a subject RC (274 ms), $t_1(39) = 6.22$, $p < .001$; $t_2(35) = 5.65$, $p < .001$.

Regression-path durations did not differ significantly for nouns embedded in object RCs (439 ms), as compared with subject RCs (453 ms), $t_1(39) = 0.57$, $p > .57$; $t_2(35) = 0.60$, $p > .55$, but were longer for verbs embedded in object RCs (513 ms) than for verbs embedded in subject RCs (408 ms), $t_1(39) = 4.37$, p < .001; $t_2(35) = 5.53$, p < .001.

Word Targeting

Our critical measures of word-targeting were *first-pass skipping rates* (i.e., the proportion of trials on which a word does not receive a first-pass fixation) and *first-pass regression rates* (i.e., the proportion of trials on which first-pass fixations on a word are followed by a regressive saccade rather than a progressive saccade). Table B1 shows skipping rates and first-pass regression rates as a function of condition for each word from the complementizer through the main-clause verb.

The complementizer was skipped more frequently in subject RCs than object RCs, $F_1(1, 39) = 23.48$, $MSE = 0.02$, $p < .001$; $F_2(1, 35) = 20.03$, $MSE = 0.02$, $p < .001$, a pattern that is somewhat puzzling given that the two types of sentences are identical up through that word. This pattern was not observed in Experiment 2 of the current article or in Staub (2010). Regressions

Table B1
Eye-Tracking Results of Experiment 1: Targeting Measures

Frequency pattern	Example word	Skips	Regressions	Example word	Skips	Regressions
	Object RC complementizer			Subject RC complementizer		
LF–LF	*that*	.23	.19	*that*	.35	.17
HF–HF	*that*	.23	.12	*that*	.31	.09
	Object RC determiner			Subject RC embedded verb		
LF–LF	*the*	.55	.36	*bothered*	.05	.28
HF–HF	*the*	.59	.25	*bothered*	.06	.29
	Object RC embedded noun			Subject RC determiner		
LF–LF	*machinist*	.06	.34	*the*	.71	.13
HF–HF	*guest*	.08	.32	*the*	.69	.08
	Object RC embedded verb			Subject RC embedded noun		
LF–LF	*bothered*	.02	.17	*machinist*	.05	.26
HF–HF	*bothered*	.04	.20	*guest*	.06	.14
	Object RC main verb			Subject RC main verb		
LF–LF	*caused*	.09	.27	*caused*	.11	.20
HF–HF	*caused*	.09	.23	*caused*	.08	.16

Note. RC = relative clause; LF = low frequency; HF = high frequency.

(*Appendices continue*)

from the complementizer occurred more frequently when the sentence contained low-frequency nouns than when it contained high-frequency nouns, $F_1(1, 39) = 8.44$, $MSE = 0.02$, $p < .007$; $F_2(1, 35) = 4.34$, $MSE = 0.02$, $p < .05$, a pattern that likely reflects difficulty in comprehending low-frequency words.

The determiner was skipped less frequently in object RCs than in subject RCs, $F_1(1, 39) = 10.32$, $MSE = 0.07$, $p < .004$; $F_2(1, 35) = 28.77$, $MSE = 0.02$, $p < .001$. This pattern likely arises because the preceding word was more likely to be skipped for object RCs (.23) than for subject RCs (.06). Regressions from the determiner occurred more frequently in object RCs than subject RCs, $(1, 39) = 12.47$, $MSE = 0.11$, $p < .002$; $F_2(1, 35) = 30.90$, $MSE = 0.04$, $p < .001$, a pattern first reported by Staub (2010). Regressions also occurred more commonly from sentences with low-frequency nouns than from those with high-frequency nouns, a pattern that was marginally significant, $F_1(1, 39) = 3.88$, $MSE = 0.06$, $p < .06$; $F_2(1, 35) = 4.80$, $MSE = 0.04$, $p < .04$.

The embedded noun showed higher rates of regressions for object RCs than for subject RCs, $F_1(1, 39) = 23.31$, $MSE = 0.03$, $p < .001$; $F_2(1, 35) = 29.42$, $MSE = 0.02$, $p < .001$, a pattern that was again first reported by Staub (2010). In addition, regression rates were higher in the low-frequency condition than in the high-frequency condition, $F_1(1, 39) = 7.51$, $MSE = 0.02$, $p < .01$; $F_2(1, 35) = 5.45$, $MSE = 0.03$, $p < .03$, a pattern that was qualified by an interaction of frequency and RC type, $F_1(1, 39) = 4.21$, $MSE = 0.02$, $p < .05$; $F_2(1, 35) = 5.63$, $MSE = 0.02$, $p < .03$. The form of this interaction is most easily described as there being fewer than expected regressions from nouns embedded in subject RCs for the high-frequency condition. This pattern suggests no obvious interpretations. The embedded verb showed lower rates of regressions for object RCs than for subject RCs, $F_1(1, 39) = 7.51$, $MSE = 0.02$, $p < .01$; $F_2(1, 35) = 5.45$, $MSE = 0.03$, $p < .03$.

Finally, the main-clause verb showed higher rates of regressions for object RCs than for subject RCs, $F_1(1, 39) = 8.00$, $MSE = 0.02$, $p < .008$; $F_2(1, 35) = 9.62$, $MSE = 0.02$, $p < .005$, as well as a trend toward higher regression rates in the low-frequency condition than the high-frequency condition, $F_1(1, 39) = 4.11$, $MSE = 0.02$, $p < .05$; $F_2(1, 35) = 3.69$, $MSE = 0.02$, $p < .07$.

Summary and Interpretation of Word-Targeting Analyses

The word-targeting analyses show some clear patterns but also provide challenges to interpretation. As discussed in the introduction, Staub (2010) has recently shown that there are higher regression rates from the determiner and noun embedded in object RCs than for those embedded in subject RCs and has argued that those effects, together with a finding of longer regression-path durations on nouns embedded in object RCs than subject RCs, support the conclusion that processing difficulty in object RCs begins with the onset of the embedded noun phrase. The current study strongly replicates the regression rate effects reported in Staub (2010). However, it does not show Staub's finding that regression-path durations are higher for nouns embedded in object RCs rather than for nouns embedded in subject RCs; it instead shows a very small (11 ms) reversal of the difference which is

not close to being statistically reliable. Thus, the greater incidence of regressions from the article and noun embedded in an object RC, as compared with a subject RC, is not associated with an increase in processing time. Since clear effects of RC type were shown on words downstream from the embedded noun, the contribution of these elevated regression rates to the processing difference between object and subject RCs is not clear. Finally, in the current study, we found higher regression rates for verbs embedded in subject RCs than for verbs embedded in object RCs, a finding not obtained in Staub (2010). This pattern is consistent with the possibility that regression rates are determined at least in part by the visual characteristics of the immediately preceding words—a relatively long head noun followed by the short, high-frequency complementizer (*that*)—which are the same for the initial constituents of object and subject RCs.

Studies focusing on eye-movement control and word recognition during reading (e.g., Brysbaert, Drieghe, & Vitu, 2005; Vitu, O'Regan, Inhoff, & Topolski, 1995) have shown that the targeting of saccades depends heavily on oculomotor processes related to length of words. Short words may be skipped because they can be recognized in the parafovea while the preceding word is fixated (Reichle et al., 1998) and also because targeting short words for fixation generally requires short saccades, a type of saccade that is prone to overshooting errors (McConkie, Kerr, Reddix, & Zola, 1988; Pollatsek et al., 2008). Regressions may also occur because of such factors, with overshooting followed by quick corrective saccades back to the previous word (Drieghe, Rayner, & Pollatsek, 2005) and with quick corrective saccades in both forward and backward directions occurring when the initial fixation on a word does not land in the optimal viewing position (Pollatsek et al., 2006). To be sure, regressive saccades also occur because of difficulties with higher level aspects of comprehension (Rayner, 1998; Reichle, Warren, & McConnell, 2009; Warren & McConnell, 2007) so that regression rates are determined by a mixture of lower level and higher level processes. Effective procedures for decomposing that mixture have not been developed, making it difficult to determine the cause of differences in regression rates for words that vary both in the length of nearby words and in terms of the complexity of the language structures in which they appear. Experiment 2 of Staub (2010) offered a control for length of nearby words by comparing sentences with object RCs to those with complement clauses, both of which precede the critical NP with the complementizer *that*. The results showed higher regression rates for the determiner and noun in the object RC, as compared with the complement clause. However, skipping rates for the complementizer (*that*) differed significantly across the two conditions, causing concern about whether the prior lexical environments were equivalent across the RC and complement clause structures. Further, the difference in regression rates for the embedded noun were smaller between the object RC and complement clause in Experiment 2 (.23 vs. .14; see Table 4, Staub, 2010) than between the object RC and subject RC in Experiment 1 (.40 vs. .16; see Table 1, Staub, 2010) and was not accompanied by a significant elevation in regression-path reading times, though there was a trend in the expected direction ($p < .10$).

(Appendices continue)

Appendix C

Reading-Time Comparisons of Word Types in Different Positions in the RC for Experiment 2

The alternative contrasts controlling for word type showed that gaze durations were shorter for nouns in object RCs (275 ms), as compared with subject RCs (325 ms), $t_1(39) = 4.60$, $p < .001$; $t_2(35) = 6.04$, $p < .001$, whereas they were longer for verbs in object RCs (360 ms) than for verbs in subject RCs (283 ms), $t_1(39) = 6.70$, $p < .001$; $t_2(35) = 10.20$, $p < .001$.

For regression-path durations, there was no significant difference for nouns in object RCs (426 ms), as compared with subject RCs (465 ms), $t_1(39) = 1.34$, $p > .18$; $t_2(35) = 1.86$, $p > .07$, whereas times were longer for verbs in object RCs (607 ms) than for verbs in subject RCs (390 ms), $t_1(39) = 6.31$, $p < .001$; $t_2(35) = 8.42$, $p < .001$.

Word Targeting

Table C1 shows skipping rates and first-pass regression rates as a function of condition for each word from the complementizer through the main clause verb. Neither skipping rates nor regression rates on the complementizer were influenced by any of the experimental variables.

The determiner was skipped less frequently in object RCs than in subject RCs, $F_1(1, 39) = 25.12$, $MSE = 0.05$, $p < .001$; $F_2(1, 35) = 48.23$, $MSE = 0.02$, $p < .001$. As in Experiment 1, this effect is likely due in part to higher rates of skipping on the

preceding word for object RCs (.30) than for subject RCs (.08). There was a trend toward more regressions from the determiner in object RCs than for the determiner in subject RCs, but it was short of significance in the subjects analysis, $F_1(1, 39) = 1.28$, $MSE = 0.09$, $p < .27$; $F_2(1, 35) = 6.96$, $MSE = 0.05$, $p < .02$. Regressions from the determiner occurred at a higher rate in the LF–HF condition than in the HF–LF condition, $F_1(1, 39) = 5.77$, $MSE = 0.07$, $p < .03$; $F_2(1, 35) = 12.11$, $MSE = 0.05$, $p < .002$.

Regressions from the embedded noun occurred at a marginally higher rate in object RCs than in subject RCs, $F_1(1, 39) = 3.26$, $MSE = 0.02$, $p < .08$; $F_2(1, 35) = 4.01$, $MSE = 0.02$, $p < .06$. Further, there was a significant interaction between type of RC and frequency pattern, such that for object RCs regression rates were higher for the HF–LF condition (.28) than the LF–HF condition (.23), whereas for subject RCs they were lower in the HF–LF condition (.18) than in the LF–HF condition (.23), $F_1(1, 39) = 11.94$, $MSE = 0.01$, $p < .002$; $F_2(1, 35) = 3.47$, $MSE = 0.02$, $p < .08$. Contrasts showed that the difference was significant by subjects, but not by items, for object RCs, $t_1(39) = 2.15$, $p < .05$; $t_2(35) = 1.20$, $p > .23$, as well as subject RCs, $t_1(39) = 2.28$, $p < .05$; $t_2(35) = 1.47$, $p > .15$. For both the embedded verb and the main-clause verb, there were no effects on skipping or regressions that reached statistical significance.

Table C1
Eye-Tracking Results of Experiment 2: Targeting Measures

Frequency pattern	Example word	Skips	Regressions	Example word	Skips	Regressions
	Object RC complementizer			Subject RC complementizer		
LF–HF	*that*	.30	.15	*that*	.33	.15
HF–LF	*that*	.29	.14	*that*	.34	.16
	Object RC determiner			Subject RC embedded verb		
LF–HF	*the*	.54	.30	*bothered*	.07	.24
HF–LF	*the*	.53	.25	*bothered*	.08	.25
	Object RC embedded noun			Subject RC determiner		
LF–HF	*guest*	.08	.23	*the*	.68	.27
HF–LF	*machinist*	.05	.28	*the*	.73	.13
	Object RC embedded verb			Subject RC embedded noun		
LF–HF	*bothered*	.06	.26	*guest*	.05	.23
HF–LF	*bothered*	.04	.29	*machinist*	.05	.18
	Object RC main verb			Subject RC main verb		
LF–HF	*caused*	.13	.29	*caused*	.11	.27
HF–LF	*caused*	.12	.25	*caused*	.10	.23

Note. RC = relative clause; LF = low frequency; HF = high frequency.

(*Appendices continue*)

Summary and Interpretation of Word-Targeting Analyses

The word-targeting analyses showed patterns that were similar overall to the previous experiment, though the greater rates of regressive saccades from the determiner and embedded noun in object RCs, as compared with subject RCs, were only marginally significant. Although the direction of these effects is the same as that observed by Staub (2010), regression-path durations on the embedded noun did not replicate Staub's finding of greater reading times for object RCs than for subject RCs, instead showing a nonsignificant advantage of 39 ms for object RCs. The results also showed an interaction between type of RC and frequency condition on regression rates from the embedded noun, with the form of the interaction being consistent with the hypothesis that the object–subject asymmetry is reduced for the LF–HF condition, as compared with the HF–LF condition. This pattern is consistent with the critical reading-time results and with the comprehension-question accuracy results. Although this regression-rate pattern is consistent with the reading time and accuracy measures, its appearance on the embedded noun is not expected based on the operation of memory-retrieval mechanisms.

Taken together, the targeting results of the two experiments suggest a mixed evaluation of the contention that differences in regression rates from determiners and embedded nouns in object RCs and subject RCs indicate a difficulty in interpreting object RCs that emerges very early in the sentence. The results provide overall corroboration of Staub's (2010) finding of higher regression rates from the determiner and noun in object RCs, as compared with subject RCs. However, as noted previously, regression rates are the product of lower level oculomotor processes related to the accuracy of targeting saccades and to higher level language-interpretation processes. Stimulus factors such as length of preceding words differ for the determiner and noun in object and subject RCs, making it uncertain how to interpret the differences in regression rates. In Staub's Experiment 1, these higher regression rates, which result from fixation patterns on a fraction of the trials, were accompanied by corresponding differences in regression-path duration, which is calculated across all of the trials. This reading-time pattern bolsters the interpretation of the differences in regression rate as reflecting a general increase in the difficulty of interpreting the determiner and noun in an object RC, as compared with a subject RC. Neither of our experiments shows this pattern for regression-path durations, with both showing trends in the opposite direction (an effect that is strongly significant in gaze duration both in our data and in Staub's).