COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science (2013) 1–17

ISSN: 0364-0213 print / 1551-6709 online
DOI: 10.1111/cogs.12021

Direct Evidence of Memory Retrieval as a Source of Difficulty in Non-Local Dependencies in Language

Evelina Fedorenko, Rebecca Woodbury, Edward Gibson

Department of Brain & Cognitive Sciences, Massachusetts Institute of Technology
Received 30 January 2010; received in revised form 11 June 2012; accepted 11 June 2012

Abstract

Linguistic dependencies between non-adjacent words have been shown to cause comprehension difficulty, compared with local dependencies. According to one class of sentence comprehension accounts, non-local dependencies are difficult because they require the retrieval of the first dependent from memory when the second dependent is encountered. According to these memory-based accounts, making the first dependent accessible at the time when the second dependent is encountered should help alleviate the difficulty associated with the processing of non-local dependencies. In a dual-task paradigm, participants read sentences that did or did not contain a non-local dependency (i.e., object- and subject-extracted cleft constructions) while simultaneously remembering a word. The memory task was aimed at making the word held in memory accessible throughout the sentence. In an object-extracted cleft (e.g., *It was Ellen whom John consulted…*), the object (*Ellen*) must be retrieved from memory when *consulted* is encountered. In the critical manipulation, the memory word was identical to the verb's object (*ELLEN*). In these conditions, the extraction effect was reduced in the comprehension accuracy data and eliminated in the reading time data. These results add to the body of evidence supporting memory-based accounts of syntactic complexity.

Keywords: Sentence comprehension; Syntactic complexity; Working memory; Non-local dependencies

1. Introduction

Interpreting syntactic dependencies among words is a critical component of language understanding. Although most such dependencies in human languages are between immediately adjacent elements (e.g., Collins, 1996; Ferrer i Cancho, 2006; Gildea & Temperley, 2009; Hawkins, 1994; Park & Levy, 2009; Temperley, 2007), in some cases words that need to be interpreted as dependent on one another are separated by other words or clauses. For example, to correctly interpret a sentence like (1), a dependency relationship

Correspondence should be sent to Evelina Fedorenko, Massachusetts Institute of Technology, 43 Vassar Street, 46-3037G, Cambridge, MA 02139. E-mail: evelina9@mit.edu

needs to be established between *girl* and *was*, in spite of the presence of a plausible subject of *was* that is immediately adjacent to it (i.e., *boy*).

(1) The girl who kissed the boy was attractive.

Dependencies between words that are not adjacent to one another have been shown to cause processing difficulty, compared with cases where the dependencies are local, manifesting in slower online processing times and lower performance in the offline tasks that require reliance on the representation of the sentence's content in several languages across many paradigms (e.g., *English*: Van Dyke & Lewis, 2003; Gibson, 1998, 2000; Gordon, Hendrick, & Johnson, 2001; Grodner & Gibson, 2005; King & Just, 1991; Lewis, Vasishth, & Van Dyke, 2006; McElree, Foraker, & Dyer, 2003; Wanner & Maratsos, 1978; *French*: Baudiffier, Caplan, Gaonach, & Chesnet, 2011; Holmes & O'Regan, 1981; *German*: Mecklinger, Schriefers, Steinhauer, & Friederici, 1995; Schriefers, Friederici, & Kühn, 1995; Vasishth & Drenhaus, 2011; but see Konieczny, 2000; Konieczny & Döring, 2003; Levy & Keller, 2012; *Dutch*: Frazier, 1987; Mak, Vonk, & Schriefers, 2002, 2006; *Japanese*: Ishizuka, Nakatani, & Gibson, 2003; Miyamoto & Nakamura, 2003; Ueno & Garnsey, 2008; *Korean*: Kwon, Polinsky, & Kluender, 2006; Kwon, Gordon, Lee, & Kluender, 2010; O'Grady, Lee, & Choo, 2003; *Russian*: Levy, Fedorenko, & Gibson, in press; cf. *Basque*: Carreiras, Duñabeitia, Vergara, de la Cruz-Pavía, & Laka, 2010; *Hindi:* Vasishth, & Lewis, 2006; *Chinese*: Gibson & Wu, 2012; Hsiao & Gibson, 2003; but see Lin & Bever, 2006). However, the source of this difficulty is still under debate. According to one class of accounts (memory-based accounts), non-local dependencies are difficult to process because they require the retrieval of the first dependent from memory when the second dependent is encountered. For example, in (1), *girl* has to be retrieved from memory when *was* is encountered.

Previous studies evaluating the predictions of memory-based accounts have typically relied on manipulations designed to make the retrieval of the non-local dependent more difficult, by increasing the distance between the two dependents, making the intervening words similar to the to-be-retrieved word, or including a secondary task designed to interfere with the retrieval of the target word (e.g., Van Dyke & Lewis, 2003; Van Dyke & McElree, 2007; Fedorenko, Gibson, & Rohde, 2006; Gordon, Hendrick, & Levine, 2002; Gordon et al., 2001; Grodner & Gibson, 2005). Greater processing difficulty—with an increase in the difficulty of the retrieval operation—has been interpreted as evidence for memory-based accounts. However, memory-based accounts make another prediction: Making the retrieval of the non-local dependent *easier* at the point of dependency formation should reduce the difficulty associated with non-local dependencies. A couple of recent findings in the literature suggest that this might be the case. For example, Hofmeister (2011) demonstrated that increasing the representational richness of the to-be-retrieved word (e.g., by adding modifying adjectives or prepositional phrases) facilitates retrieval. Relatedly, Vasishth, Shaher, Logacev, Engelmann, and Srinivasan (in press) have observed that nouns in more prominent syntactic positions (clefted nouns) are easier to retrieve than their non-clefted counterparts in Hindi. The current experiment provides a more direct evaluation of this prediction of memory-based accounts.

2. Experiment

One well-studied contrast between local and non-local dependency structures involves subject- and object-extracted relative clauses or clefts, as in (2):

(2) a. It was John who consulted Ellen in the library.
  b. It was Ellen who John consulted in the library.

The object-extracted cleft (2b) contains a non-local dependency between the pronoun *who* (referring to *Ellen*) and the verb *consulted.* In contrast, in the subject-extracted cleft (2a), *consulted* is immediately adjacent to both its subject *who* (referring to *John*) and its object *Ellen*.

Sentences containing subject- versus object-extractions are commonly used for investigating structural dependencies because they differ only in the positions of the critical words, so that the critical comparisons are performed on similar (often identical) words that are located in similar positions in the sentence. Furthermore, it is easy to match such materials for plausibility, either by fixing the meaning across the two structures as in (2), or by including two versions of each structure (e.g., using *It was Ellen who consulted John in the library* and *It was John who Ellen consulted in the library* in addition to the versions in (2)). We therefore chose to use subject- and object-extracted cleft structures in the current experiment.

To achieve the desired manipulation—facilitating the retrieval of the non-local dependent at the point of dependency formation—we adapted a dual-task paradigm in which participants see a word or a set of words that they are instructed to remember, then read a sentence word-by-word, and then report the word(s) held in memory and answer a question about the sentence (e.g., Fedorenko et al., 2006; Gordon et al., 2002). We made two modifications to the earlier versions of this paradigm. First, we made the memory task easy by having participants keep at most one word in memory. And second, in the critical condition, we made the word held in memory identical to the object noun. If the difficulty in the object-extracted structures is due to the need to retrieve the object noun from memory, then making the object noun salient throughout the sentence (via the memory task)—thereby eliminating the need to retrieve it from memory at the verb—should partially or completely alleviate the difficulty typically associated with the processing of object-extracted structures.

3. Methods

3.1. Participants

Sixty native English speakers between 18 and 40 years old—students at MIT and members of the surrounding community—participated in the study. All participants were paid for their participation and were naive to the purposes of the study.

3.2. Design and materials

The experiment manipulated the cleft structure (subject-extracted, object-extracted) and the presence and kind of the word held in memory (identical to the object of the embedded verb, a control noun that did not appear in the sentence, identical to the subject of the embedded verb, or no word to remember) in a 2 × 4 design, as illustrated in (3). The critical prediction concerns the *object-memory-word* object-extracted condition, where the object noun (*Ellen*) should no longer be difficult to retrieve at the point of processing the verb because the object is made salient throughout the sentence and should therefore be more easily accessible when the verb-object dependency is formed. This manipulation should therefore lead to faster processing times and a more robust memory representation of the sentence, compared with the *no-memory-word* condition.

The *control-memory-word* conditions were included to determine the effect on the sentence-reading task of the secondary (memory) task, thus providing an additional baseline for the critical *object-memory-word* conditions. This manipulation should make the object-extracted condition (and possibly the subject-extracted condition) more difficult than the *no-memory-word* conditions (Fedorenko et al., 2006; Gordon et al., 2002).

Finally, the *subject-memory-word* conditions were included to evaluate an alternative hypothesis about why processing a sentence like *It was Ellen who John consulted in the library* might be facilitated when the word *ELLEN* is held in memory. In particular, *Ellen* is the topic of the sentence, and having the memory word match the topic may facilitate the topic's encoding leading to a stronger memory trace and a facilitatory effect on the processing of the rest of the sentence. According to this hypothesis, in addition to the facilitation in the object-memory-word object-extracted condition, we should see faster processing times and higher comprehension accuracies in the subject-memory-word subject-extracted condition compared with the subject-extracted condition where no word is held in memory.

Note that both (a) the critical hypothesis about the facilitated retrieval of the object noun from memory in the object-memory-word object-extracted condition, and (b) the alternative hypothesis about the general facilitation resulting from a more robust encoding of the sentence topic (which predicts facilitation in both the object-memory-word object-extracted condition and in the subject-memory-word subject-extracted condition), appeal to memory processes. However, whereas the latter focuses on the encoding of the clefted noun, the former focuses on its retrieval at the point of dependency formation (see e.g., Hofmeister, 2011; Lewis et al., 2006, for a discussion of encoding/retrieval processes in sentence comprehension). It is also worth noting that the two hypotheses are not mutually exclusive.

(3)

Condition	Memory Word	Sentence
No memory word/Subj-extracted	X	It was John who consulted Ellen in the library.
No memory word/Obj-extracted	X	It was Ellen who John consulted in the library.
Obj. mem. word/Subj-extracted	ELLEN	It was John who consulted Ellen in the library.
Obj. mem. word/Obj-extracted	ELLEN	It was Ellen who John consulted in the library.
Ctrl mem. word/Subj-extracted	STEVE	It was John who consulted Ellen in the library.
Ctrl mem. word/Obj-extracted	STEVE	It was Ellen who John consulted in the library.
Subj. mem. word/Subj-extracted	JOHN	It was John who consulted Ellen in the library.
Subj. mem. word/Obj-extracted	JOHN	It was Ellen who John consulted in the library.

Forty sets of sentences were constructed (see Appendix). Each participant saw only one version of each item, following a Latin-Square design. Names were used instead of occupation nouns—often used in linguistic materials—for three reasons. First, using names largely eliminates the issues with potential differences in plausibility between the subject- and object-extracted conditions. In particular, in the sentences where a name is used in both the subject and the object position the only cue to the dependency structure is the syntax of the sentence (i.e., the positions of the nouns relative to the verb and to each other in this case). Second, somewhat relatedly, if we used occupation nouns, we would have to control for the potential semantic relationships between the nouns (both sentence and memory nouns) and the verb. The effect of the memory word on the processing of the sentence may therefore be due not only to the relationship between the memory word and the sentence nouns but also to the relationship between the memory word and the sentence verb. Using names allowed us to examine retrieval difficulty as a function of how accessible in memory the target noun is, without worrying about the verb's semantics. Finally, using names reduces the potential variability that may be associated with the familiarity/frequency of different common nouns (such as occupations), allowing us to focus on the manipulation of interest, that is, on the dependency structure of the sentences. Because we used names, we had to use the cleft construction instead of the relative clause construction because only the former allows the use of names in the head noun position.

In constructing the materials, we had to make a choice between (a) keeping the words in the critical region (the embedded noun and the verb) identical across the subject- and the object-extracted conditions, and (b) keeping the memory word identical—within each memory-word condition—across the subject and the object-extracted conditions. We chose the latter because we thought it was more important to compare the effects of the same memory word on the processing of sentences that have the same meaning but differ in structure. To control for potential issues of comparing across different words in the subject- and object-extracted conditions, each name appeared once in each syntactic position across the item set. For example, in addition to the item in (2) where *John* is used in the subject position, there was another item where *John* is used in the object position. Similarly, there was another item where *Ellen* is used in the subject position.

Verbs were chosen such that a male or a female was similarly likely to be the agent or the patient. The gender of the names was counterbalanced (10 items used two male names; 10—two female names; 10—male subject and female object; and 10—female subject and male object). For the sentences where both nouns were either male or female, the control memory word matched the sentence nouns in gender. For the sentences where one name was male and one female, the control memory word was male half of the time.

To assess the robustness of the memory representations of the sentence content, a yes/no comprehension question about the propositional content of the sentence was asked at the end of each trial. Two question types were used, each requiring the understanding of the dependency structure of the sentence: (a) a question asking only about the dependency structure (e.g., *Did John consult Ellen? Yes/Did Ellen consult John? No,* for the item in (2)); and (b) a question asking about both the dependency structure and the sentence ending (e.g., *Did John consult Ellen in the library? Yes/Did Ellen consult John in the classroom? No*). Within an item, the same question was asked for all the conditions. Question types varied across items and the number of "yes" and "no" questions was balanced.

The experiment also included 80 filler sentences containing names, which were similar to the critical sentences in length and complexity. In half of the fillers, the memory word was a name that appeared in the sentence, in a quarter of the fillers, the memory word was a name that did not appear in the sentence, and in the remaining quarter there was no memory word to match the distribution of memory words in the critical sentences. (The fillers are available from the authors upon request.)

Trials within each experimental list were randomized separately for each participant.

3.3. Procedure

The sentence-reading task used self-paced moving-window word-by-word reading (Just, Carpenter, & Woolley, 1982). The experiment was run using the Linger 2.85 software developed by Doug Rohde (available at http://tedlab.mit.edu/~dr/Linger/).

Each trial began with a word or an "X" appearing on the screen in capital letters for 600 ms. An "X" was used for the no-memory-word trials, to match the structure of the trials in the other conditions. Participants were instructed to remember the word. A blank screen was then presented for 500 ms, which was followed by a series of dashes marking the length and position of the words in the sentence. Participants pressed the spacebar to reveal each word of the sentence. As each new word appeared, the preceding word was converted to dashes again. The amount of time participants spent reading each word was recorded as the time between key presses. After the sentence, a box appeared, and participants were instructed to type in the word from the memory task or an "X" (for the no-memory-word trials). Then a comprehension question appeared on the screen. Participants pressed one of two keys to respond "yes"/"no". At the end of each trial, participants were shown the percentage of correct responses in the experiment so far, across the memory and the comprehension tasks. They were told to take low percentages as an indication to be more careful.

Before the experiment, a short list of practice items was presented to familiarize participants with the task. The experiment took approximately 40 min.

4. Results

Before conducting detailed analyses, we examined individual subjects' accuracies on the memory task and the comprehension question task. We removed four subjects with accuracies lower than 67% on either the memory task or the comprehension question task. For the remaining 56 subjects, the across-conditions mean was 96.7% for the memory task and 90.7% for the comprehension question task.

Analyses reported here were conducted with the lme4 package (Bates, Maechler, & Dai, 2008) for the statistical language R (R Core Development Team, 2008). The data from the memory task and the comprehension task were analyzed with logistic regressions. The reading time data were analyzed with a linear mixed-effects regression. Recent results have shown that including only random intercepts in linear mixed-effects regressions can be anti-conservative (e.g., Barr, Levy, Scheepers, & Tily, in press), so we also include random slopes for participants in modeling the reading time data. (Further inclusion of random slopes for items led to a lack of model convergence.) Significance (p) values were estimated from (a) the t values that were obtained from the lmer function; and (b) conservative estimates of the number of degrees of freedom in the model. The estimates of the number of degrees of freedom in the model consisted of the number of observations minus the number of intercepts fit in the model (the number of participants + the number of items = 56 + 40 = 96) and the number of slopes being fit in the model (the number of participants = 56).

4.1. *Memory task performance*

As intended, the memory task was easy, as evidenced by close-to-ceiling performance (Table 1).[1] We performed a logistic regression with two sum-coded factors: extraction (subject-, object-) and type of memory word (object memory word, control memory word, subject memory word). The control-memory-word object-extracted condition was chosen as the baseline. The results of the model are summarized in Table 2. The extraction manipulation did not have a significant effect on accuracy. For the memory word manipulation, the subject-memory-word conditions were not reliably different from the baseline, but the object-memory-word conditions were reliably better. No interactions were observed. The small but reliable improvement in the object-memory-word conditions, compared with the control-memory-word conditions, is plausibly due to the overall better performance on the object-memory-word conditions, as discussed in the sections below.

4.2. *Comprehension accuracies*

Accuracies for the eight conditions are shown in Fig. 1. We performed a logistic regression with two sum-coded factors: extraction (subject-, object-) and type of memory word (no memory word, object memory word, control memory word, subject memory word). The no-memory-word object-extracted condition was chosen as the baseline. The

Table 1
Memory task performance in the six conditions (standard errors of the by-participants' mean in parentheses)

	Memory Word		
Structure	Object	Subject	Control
Subject-extracted	.986 (.01)	.982 (.01)	.961 (.01)
Object-extracted	.989 (.01)	.968 (.01)	.918 (.02)

Table 2
Results of the logistic regression on the memory task data

Coefficients	Estimate	Std. Error	*z* Value	Pr(>|*z*|)
(Intercept)	3.62909	0.17087	21.239	<2e−16***
Extraction.subj	0.18242	0.17087	1.068	0.28568
MemWord.object	0.75067	0.28002	2.681	0.00734*
MemWord.subject	0.07703	0.23608	0.326	0.74421
Extraction.subj:MemWord.object	−0.32807	0.28002	−1.172	0.24135
Extraction.subj:MemWord.subject	0.11879	0.23608	0.503	0.61482

Note. Significance codes: 0 "***" 0.001 "**" 0.01 "*" 0.05 "††" 0.1 "†" 1.

results of the model are summarized in Table 3. The extraction manipulation had a highly reliable effect on accuracy. For the memory word manipulation, we observed two main effects: The object-memory-word conditions were reliably more accurate, and the control-memory-word conditions were reliably less accurate than the baseline (no-memory-word conditions). Lower accuracies in the control-memory-word conditions are consistent with previous findings (e.g., Fedorenko et al., 2006; Gordon et al., 2002). Furthermore, we observed a significant interaction, such that the extraction effect was smaller for the object-memory-word conditions compared with the no-memory-word conditions.

4.3. Reading times

Before performing linear mixed-effects regression analyses on the reading time data, we removed raw reading times that were longer than 5,000 ms and then transformed the raw reading times into residual reading times (Ferreira & Clifton, 1986; see Trueswell, Tanenhaus, & Garnsey, 1994; for discussion). First, a regression equation predicting reading times from region length was derived for each participant, using all filler and target items. Then, for each region, the reading time predicted by the participant's regression equation was subtracted from the actual measured reading time to obtain a residual reading time. We analyzed residual reading times because (a) they adjust for differences in region lengths and overall differences in participants' reading rates; and (b) they are more normally distributed than raw reading times (normally distributed data are a requirement for regression analyses, cf. Jaeger, 2008). Residual reading times greater than three stan-

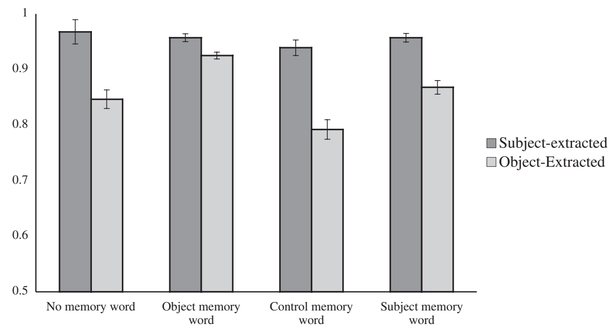

Fig. 1. Comprehension question accuracies in the eight conditions. Error bars represent standard errors of the by-participants' mean.

Table 3
Results of the logistic regression on the comprehension question data

Coefficients	Estimate	Std. Error	z Value	Pr(>|z|)
(Intercept)	2.473440	0.086913	28.459	<2e−16***
Extraction.subj	0.613683	0.086913	7.061	1.65e−12***
MemWord.subj	0.020672	0.149431	0.138	0.88997
MemWord.object	0.335753	0.157708	2.129	0.03326*
MemWord.control	−0.437014	0.134554	−3.248	0.00116**
Extraction.subj:MemWord.subj	0.001714	0.149431	−0.011	0.99085
Extraction.subj:MemWord.object	0.316795	0.157708	−2.009	0.04456*
Extraction.subj:MemWord.control	0.085023	0.134554	0.632	0.52746

Note. Significance codes: 0 "***" 0.001 "**" 0.01 "*" 0.05 "††" 0.1 "†" 1.

dard deviations away from the mean for a position within condition were removed from the analyses. These procedures resulted in the removal of 3.2% of the data.

We defined the critical region as the portion of the sentence where the dependency structure is manipulated (*consulted Ellen/John consulted* in (2)), as is commonly done in the literature. Although the critical effects are predicted to occur at the verb (*consulted*), this word alone is not directly comparable across the subject- and the object-extracted conditions because (a) the verb occurs in different sentence positions and position within the sentence is known to affect reading times, and even more important, (b) in the subject-extracted conditions at the point where the verb is processed the object noun has not yet been encountered and therefore the memory word manipulation in the object-memory-word and control-memory-word conditions has not yet had a chance to affect sentence processing. Fig. 2 presents the RTs for the critical region across the eight conditions (see Appendix for the sentence word-by-word RT graphs).

We performed a linear mixed-effects regression with two sum-coded factors: extraction (subject-, object-) and type of memory word (no memory word, object memory word, control memory word, subject memory word). As in the analysis of comprehension accuracies, the no-memory-word object-extracted condition was chosen as the baseline.

To remove the potential contribution of the facilitated encoding of the clefted noun, reading times on the region preceding the critical region (i.e., *who*) were included as an additional predictor in the model, per a reviewer's suggestion. The results of the model are summarized in Table 4. The extraction manipulation (as well as the reading times on the preceding region) had highly reliable effects on reading times (both *p*s < .005). No main effects were observed for the memory word manipulation. However, a significant interaction between extraction and memory word manipulations was observed, such that the extraction effect was smaller for the object-memory-word conditions compared with the no-memory-word conditions ($p < .05$). There was additionally a marginal interaction such that the extraction effect was larger for the subject-memory-word conditions compared with the no-memory-word conditions ($p = .084$).

5. Discussion

We here tested a prediction of memory-based accounts of syntactic complexity. According to these accounts, non-local dependencies are more costly than local dependencies because in the former the first element of the dependency must be retrieved from memory when the second element of the dependency is encountered (unlike local dependencies, where the first element of the dependency is still highly active when the second element is encountered). These accounts therefore predict that making the to-be-retrieved element highly active in memory at the point when the dependency is formed should reduce the difficulty typically associated with non-local dependencies.

We used a dual-task paradigm in an attempt to facilitate the retrieval of the object noun from memory in object-extracted cleft constructions by making the object noun highly accessible at the point of the dependency formation. For example, participants were holding in memory the word *ELLEN* while processing the sentence *It was Ellen who John consulted at the library*. When *consulted* is encountered and the dependency between *consulted* and its object (*Ellen*) needs to be established, the object noun is

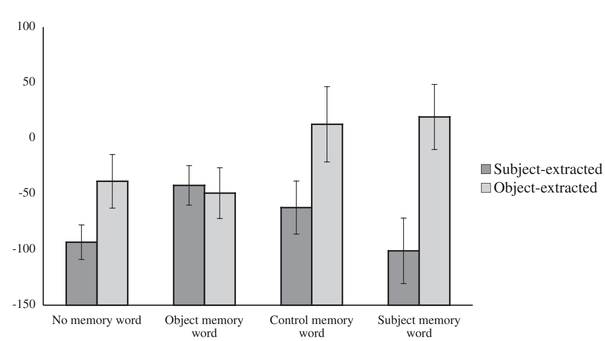

Fig. 2. Reading times at the critical region in the eight conditions. Error bars represent standard errors of the by-participants' mean.

Table 4
Results of the linear mixed-effects regression on the reading time data

Coefficients	Estimate	Std. Error	*t* Value
(Intercept)	−29.71625	19.24355	−1.544
PrecRegion.RT	0.69848	0.05971	11.697
Extraction.subj	−29.36982	9.09843	−3.228
Related.sumCoding	7.63745	10.88957	0.701
Related.sumCoding2	−2.02203	10.93088	−0.185
Related.sumCoding3	12.48072	12.34064	1.011
Extraction.subj:MemWord.subj	−25.52178	14.43754	−1.768
Extraction.subj:MemWord.object	22.00434	10.21489	2.154
Extraction.subj:MemWord.control	−3.47168	11.92366	−0.291

already highly active in memory. Consistent with the prediction of memory-based accounts, holding in memory the object noun during the processing of object-extracted conditions led to a robust memory representation of the dependency structure: Comprehension question accuracies were higher in this condition than in the other three object-extracted conditions (even in the no-memory-word condition, which had no memory-task demands). A similar pattern was observed in the reading times at the critical region: Reading times in the object-memory-word condition were numerically faster than in the other three object-extracted conditions, and the extraction effect was eliminated in the object-memory-word conditions.

What is the nature of the facilitation effect observed in the object-memory-word object-extracted condition? As discussed in Methods, an alternative to the retrieval facilitation hypothesis is a hypothesis whereby a match between the memory word and the topic of the sentence facilitates the encoding of the topic noun, leading to a facilitatory effect on the processing of the sentence. This hypothesis predicts faster processing times and higher comprehension accuracies not only in the object-memory-word object-extracted condition but also in the subject-memory-word subject-extracted condition.

The reading time data patterns provide some support for the encoding-based hypothesis. In particular, at the clefted noun region, the subject-memory-word condition is numerically the fastest of the four subject-extracted conditions, and the object-memory-word condition is the fastest of the four object-extracted conditions (see Appendix). Furthermore, at the critical region, the conditions where the memory word mismatches the topic are slower than those where the memory word matches the topic (or where no word is held in memory). Importantly, however, in our analysis of the reading time data we included reading times on the preceding region (*who*) as a predictor in the model, and the critical effects remained significant. Moreover, the comprehension accuracy data pattern is consistent with the retrieval-based, but not the encoding-based hypothesis. In particular, the subject-memory-word subject-extracted condition is less, not more, accurate than the no-memory-word subject-extracted condition, and the subject-memory-word object-extracted condition is more, not less, accurate than the no-memory word object-extracted condition. As a result, we tentatively conclude that the effect observed in the object-memory-word object-extracted condition is largely due to facilitated retrieval of the object noun, but we leave open the possibility of enhanced encoding playing some role too. As noted above, these hypotheses are not mutually exclusive.

Until now we have been focusing on memory-based accounts of syntactic complexity. According to an alternative class of accounts, the difficulty in non-local dependencies has to do with the lower frequency of these structures in the input, compared with local-dependency structures (e.g., Gennari & MacDonald, 2008; Hale, 2001; Keenan & Comrie, 1977; Levy, 2008; Reali & Christiansen, 2007). Whereas much existing evidence concerning subject- and object-extracted structures is consistent with both memory-based and experience-based accounts, some evidence is better explained by one or the other. Experience-based accounts are better able to explain the effects of position-dependent lexico-semantic properties of the noun phrases on processing difficulty. For example, the difficulty of processing object-extractions is affected by whether the subject or the object is animate versus inanimate (e.g., Traxler, Morris, & Seely, 2002; but see Gibson, Tily, & Fedorenko, in press; Tily, Fedorenko, & Gibson, unpublished data), or whether the subject or the object is a pronoun (e.g., Reali & Christiansen, 2007; Warren & Gibson, 2002): The less frequent configurations lead to more processing difficulty.

On the other hand, memory-based accounts are better able to explain (1) the locus of the processing difficulty in object-extracted structures, and (2) similarity-based interference effects in normal reading and in dual-task paradigms. First, most of the difficulty in object-extracted structures occurs at the verb (cf. Staub, 2010). This is straightforwardly explained by memory-based accounts because the verb is where the dependency is formed and where therefore the retrieval of the object noun from memory takes place. In contrast, experience-based accounts predict processing difficulty to occur as soon as the comprehender can determine that an object-extracted structure has been encountered, which is at the embedded subject. However, little, if any, processing difficulty is experienced during the embedded subject or during the material that can be inserted between the embedded subject and embedded verb (Grodner & Gibson, 2005; Gibson et al., in press; cf. Staub, 2010). And second, the difficulty of processing object-extractions increases when the subject and the object are syntactically and semantically similar (Gordon, Hendrick, & Johnson, 2004; Gordon et al., 2001), or when participants have to hold in memory words that are similar to the nouns in the sentence (Fedorenko et al., 2006; Gordon et al., 2002) or that compete with the object noun for being interpreted as the object of the verb (Van Dyke & McElree, 2007). These results are predicted by memory-based accounts, but not by experience-based accounts. The results reported here provide new direct evidence in support of memory-based accounts, which cannot be straightforwardly explained by experience-based accounts. As many researchers now concur, features of both memory-based and experience-based accounts are necessary to explain the wealth of the available evidence (e.g., Boston, Hale, Vasishth, & Kliegl, 2011; Demberg & Keller, 2008; Gibson et al., in press; Lewis et al., 2006).

Acknowledgments

We thank the audience at CUNY 2009 (UC Davis), TedLab members, Shravan Vasishth, and two anonymous reviewers for helpful comments on this work. We are also grateful to Peter Graff, Steve T. Piantadosi, and Hal Tily for advice on the analyses. This research was supported by grant 0844472 from the Linguistics Program at the National Science Foundation to EG.

Note

1. Deviations from the target by one letter were not counted as errors.

References

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (in press). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*

Bates, D., Maechler, M., & Dai, B. (2008). lme4: Linear mixed-effects models using S4 classes. Available at: http://lme4.r-forge.r-project.org/. Accessed June 1, 2012.

Baudiffier, V., Caplan, D., Gaonach, D., & Chesnet, D. (2011). The effect of noun animacy on the processing of unambiguous sentences: Evidence from French relative clauses. *Quarterly Journal of Experimental Psychology*, *64*(10), 1896–1905.

Boston, M. F., Hale, J. T., Vasishth, S., & Kliegl, R. (2011). Parallel processing and sentence comprehension difficulty. *Language and Cognitive Processes*, *26*(3), 301–349.

Carreiras, M., Duñabeitia, J. A., Vergara, M., de la Cruz-Pavía, I., & Laka, I. (2010). Subject relative clauses are not universally easier to process: Evidence from Basque. *Cognition*, *115*, 19–92.

Collins, M. (1996). A new statistical parser based on bigram lexical dependencies. Proceedings of the 34th Annual Meeting of ACL, Santa Cruz.

Demberg, V., & Keller, F. (2008). Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition*, *109*(2), 193–210.

Fedorenko, E., Gibson, E., & Rohde, D. (2006). The nature of working memory capacity in sentence comprehension: Evidence against domain-specific working memory resources. *Journal of Memory and Language*, *54*, 541–553.

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

Ferrer i Cancho, R. (2006). Why do syntactic links not cross? *Europhysics Letters*, *76*, 1228–1235.

Frazier, L. (1987). Syntactic processing evidence from Dutch. *Natural Language and Linguistic Theory*, *5*, 519–559.

Gennari, S., & MacDonald, M. (2008). Semantic indeterminacy in object relative clauses. *Journal of Memory and Language*, *58*, 161–187.

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

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*. Cambridge, MA: MIT Press.

Gibson, E., Tily, H., & Fedorenko, E. (in press). The processing complexity of English relative clauses.

Gibson, E., & Wu, I. (2012). Processing Chinese relative clauses in context. *Language and Cognitive Processes*.

Gildea, D., & Temperley, D. (2009). Do grammars minimize dependency length? *Cognitive Science*, *34*, 286–310.

Gordon, P. C., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *27*(6), 1411–1423.

Gordon, P., Hendrick, R., & Johnson, M. (2004). Effects of noun phrase type on sentence complexity. *Journal of Memory and Language*, *51*, 97–114.

Gordon, P. C., Hendrick, R., & Levine, W. H. (2002). Memory load interference in syntactic processing. *Psychological Science*, *13*, 425–430.

Grodner, D., & Gibson, E. (2005). Consequences of the serial nature of linguistic input. *Cognitive Science*, *29*(2), 261–290.

Hale, J. (2001). A probabilistic Earley parser as a psycholinguistic model. *Proceedings of NAACL*, *2*, 159–166.

Hawkins, J. (1994). *A performance theory of order and constituency. Cambridge Studies in Linguistics*. Cambridge, England: Cambridge University Press.

Hofmeister, P. (2011). Representational complexity and memory retrieval in language comprehension. *Language and Cognitive Processes*, *26*(3), 376–405.

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.

Hsiao, F., & Gibson, E. (2003). Processing relative clauses in Chinese. *Cognition*, *90*, 3–27.

Ishizuka, T., Nakatani, K., & Gibson, E. (2003). *Relative clause extraction complexity in Japanese*. Poster presented at the 16th annual CUNY conference on human sentence processing. Cambridge, MA: Massachusetts Institute of Technology.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models. *Journal of Memory and Language*, *59*, 434–446.

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processing in reading comprehension. *Journal of Experimental Psychology: General*, *111*, 228–238.

Keenan, E., & Comrie, B. (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry*, *8*, 63–99.

King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language*, *30*, 580–602.

Konieczny, L. (2000). Locality and parsing complexity. *Journal of Psycholinguistic Research*, *29-6*, 627–645.

Konieczny, L., & Döring, P. (2003). Anticipation of clause-final heads: Evidence from eye-tracking and SRNs. In Proceedings of ICCS/ASCS.

Kwon, N., Polinsky, M., & Kluender, R. (2006). Subject preference in Korean. In D. Baumer, D. Montero & M. Scanlon (Eds.), *Proceedings of the 25th West Coast Conference on Formal Linguistics* (WCCFL 25), (pp. 1–14). Somerville, MA: Cascadilla Press.

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

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*(3), 1126–1177.

Levy, R., Fedorenko, E., & Gibson, E. (in press). The syntactic complexity of Russian relative clauses. *Journal of Memory and Language*

Levy, R., & Keller, F. (2013). Expectation and locality effects in German verb-final structures. *Journal of Memory and Language*, *68*, 199–222.

Lewis, R., Vasishth, S., & Van Dyke, J. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences*, *10*(10), 447–454.

Lin, C. J. C., & Bever, T. G. (2006). Subject preference in the processing of relative clauses in Chinese. In D. Baumer, D. Montero & M. Scanlon (Eds.), *Proceedings of the 25th West Coast conference on formal linguistics (WCCFL 25)* (pp. 254–260). Sommerville, MA: Cascadilla Press.

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

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

McElree, B., Foraker, S., & Dyer, L. (2003). Memory structures that subserve sentence comprehension. *Journal of Memory and Language*, *48*, 67–91.

Mecklinger, K., Schriefers, H., Steinhauer, K., & Friederici, A. (1995). Processing relative clauses varying on syntactic and semantic dimensions. *Memory and Cognition*, *23*, 477–494.

Miyamoto, E., & Nakamura, M. (2003). Subject/object asymmetries in the processing of relative clauses in Japanese. In G. Garding & M. Tsujimura (Eds.), *Proceedings of the 22nd West Coast conference on formal linguistics* (pp. 342–355). Somerville, MA: Cascadilla Press.

O'Grady, W., Lee, M., & Choo, M. (2003). Asubject-object asymmetry in the acquisition of relative clauses in Korean as a second language. *Studies in Second Language Acquisition*, *25*, 433–448.

Park, Y. A., & Levy, R. (2009). Minimal-length linearizations for mildly context-sensitive dependency trees. Proceedings of the North American Chapter of the Association for Computational Linguistics – Human Language Technologies (NAACL-HLT) conference.

Reali, F., & Christiansen, M. H. (2007). Processing of relative clauses is made easier by frequency of occurrence. *Journal of Memory and Language*, *53*, 1–23.

R Development Core Team (2008). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

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

Staub, A. (2010). Eye movements and processing difficulty in object relative clauses. *Cognition*, *116*, 71–86.

Temperley, D. (2007). Minimization of dependency length in written English. *Cognition*, *105*, 300–333.

Traxler, M., Morris, R., & Seely, R. (2002). Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory and Language*, *47*, 69–90.

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

Ueno, M., & Garnsey, S. (2008). An ERP study of the processing of subject and object relative clauses in Japanese. *Language and Cognitive Processes*, *23*, 646–688.

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.

Van Dyke, J. A., & McElree, B. (2007). Retrieval interference in sentence comprehension. *Journal of Memory and Language*, *55*, 157–166.

Vasishth, S., & Drenhaus, H. (2011). Locality in German. *Dialogue and Discourse*, *1*, 59–82.

Vasishth, S., & Lewis, R. (2006). Argument-head distance and processing complexity: Explaining both locality and anti-locality effects. *Language*, *82*(4), 767–794.

Vasishth, S., Shaher, R., Logacev, P., Engelmann, F., & Srinivasan, N. (in press). The role of clefting, word order and given-new ordering: Evidence from Hindi. *Journal of South Asian Linguistics*, *4*(2).

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, MA: MIT Press.

Warren, T., & Gibson, E. (2002). The influence of referential processing on sentence complexity. *Cognition*, *85*, 79–112.

Appendix: Experimental materials

We first demonstrate how the eight conditions were created using the first item. For the rest of the items we show the subject-extracted version and provide the control memory noun. The rest of the conditions can be generated as shown for item #1.

1. It was John who consulted Ellen in the library. [STEVE]

Subject-extracted conditions:
No memory word: [X] It was John who consulted Ellen in the library.
Object memory word: [ELLEN] It was John who consulted Ellen in the library.
Control memory word: [STEVE] It was John who consulted Ellen in the library.
Subject memory word: [JOHN] It was John who consulted Ellen in the library.
Object-extracted conditions:
No memory word: [X] It was Ellen who John consulted in the library.
Object memory word: [ELLEN] It was Ellen who John consulted in the library.
Control memory word: [STEVE] It was Ellen who John consulted in the library.
Subject memory word: [JOHN] It was Ellen who John consulted in the library.

2. It was Sara who distracted George at the meeting. [EMILY]
3. It was Ellen who offended Sara during the class. [JENNIFER]
4. It was George who criticized John after the talk. [MATT]
5. It was Edward who helped Caroline in the park. [ANNA]
6. It was Jennifer who congratulated Steve at the show. [JEFFREY]
7. It was Caroline who interviewed Jennifer in the office. [SOPHIE]
8. It was Steven who called Edward from the station. [JACOB]
9. It was Dan who emailed Sally about the job. [MICHAEL]
10. It was Jill who noticed Pete at the supermarket. [MEG]
11. It was Sally who applauded Jill after the debate. [RACHEL]
12. It was Pete who disliked Dan after the incident. [DAVE]
13. It was Joe who hired Laura for the position. [REBECCA]
14. It was Melissa who rewarded Jim with the bonus. [GREG]
15. It was Laura who tutored Melissa before the exam. [ABIGAIL]
16. It was Jim who followed Joe to the building. [ETHAN]
17. It was Noah who kicked Emma in the shin. [ALEX]
18. It was Olivia who neglected Aaron at the party. [HANNAH]
19. It was Emma who harassed Olivia in the bookstore. [LILY]
20. It was Aaron who misinformed Noah about the deal. [ANDREW]
21. It was Jack who misquoted Grace in the article. [ISABELLE]
22. It was Elizabeth who admired Nick for his courage. [JOSH]
23. It was Grace who warned Elizabeth about the storm. [AMELIA]
24. It was Nick who chased Jack around the lake. [OWEN]
25. It was Zachary who blamed Alyssa for the mistake. [WILLIAM]
26. It was Allison who thanked Adam for the gifts. [ABBY]
27. It was Alyssa who picked Allison for the team. [CLAIRE]

28. It was Adam who chaperoned Zachary to the dance. [IAN]
29. It was Colin who greeted Sophie at the door. [NATALIE]
30. It was Eva who punished James for his actions. [RYAN]
31. It was Sophie who attacked Eva at the conference. [SAMANTHA]
32. It was James who confused Collin during the discussion. [SETH]
33. It was Nathan who defended Lauren at the interview. [SAMUEL]
34. It was Amy who accused Tony at the hearing. [AMBER]
35. It was Lauren who provoked Amy during the argument. [LEAH]
36. It was Tony who stabbed Nathan on the street. [EVAN]
37. It was Justin who visited Erin at her house. [LUCY]
38. It was Victoria who inspired Tyler at the gallery. [BRENDAN]
39. It was Erin who surprised Victoria at the airport. [JULIA]
40. It was Tyler who described Justin to the police. [HENRY]

Residual reading times across the sentence