COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science (2014) 1–25

ISSN: 0364-0213 print / 1551-6709 online
DOI: 10.1111/cogs.12168

Syntactic Complexity Effects in Sentence Production

Gregory Scontras,[a] William Badecker,[b] Lisa Shank,[c] Eunice Lim,[d] Evelina Fedorenko[e]

[a]*Department of Psychology, Stanford University*
[b]*Cognitive Science Program, University of Arizona*
[c]*Uniformed Services University of the Health Sciences*
[d]*Department of Cognitive Science, University of California, San Diego*
[e]*Department of Psychiatry, Massachusetts General Hospital*

Received 8 January 2013; received in revised form 29 December 2013; accepted 28 January 2014

Abstract

Syntactic complexity effects have been investigated extensively with respect to comprehension (e.g., Demberg & Keller, 2008; Gibson, 1998, 2000; Gordon et al., 2001, 2004; Grodner & Gibson, 2005; King & Just, 1991; Lewis & Vasishth, 2005; Lewis et al., 2006; McElree et al., 2003; Wanner & Maratsos, 1978). According to one prominent class of accounts (experience-based accounts; e.g., Hale, 2001; Levy, 2008; Gennari & MacDonald, 2008, 2009; Wells et al., 2009), certain structures cause comprehension difficulty due to their scarcity in the language. But why are some structures less frequent than others? In two elicited-production experiments we investigated syntactic complexity effects in relative clauses (Experiment 1) and wh-questions (Experiment 2) varying in whether or not they contained non-local dependencies. In both experiments, we found reliable durational differences between subject-extracted structures (which only contain local dependencies) and object-extracted structures (which contain nonlocal dependencies): Participants took longer to begin and produce object-extractions. Furthermore, participants were more likely to be disfluent in the object-extracted constructions. These results suggest that there is a cost associated with planning and uttering the more syntactically complex, object-extracted structures, and that this cost manifests in the form of longer durations and disfluencies. Although the precise nature of this cost remains to be determined, these effects provide one plausible explanation for the relative rarity of object-extractions: They are more costly to produce.

Keywords: Syntactic complexity; Sentence production; Sentence processing; Working memory; Relative clauses; Wh-questions

Correspondence should be sent to Gregory Scontras, Department of Psychology, Stanford University, 450 Serra Mall, Stanford, CA 94305. E-mail: scontras@stanford.edu; or Evelina Fedorenko, Massachusetts General Hospital, Department of Psychiatry, Building 149, East 13th Street, Charlestown, MA 02129. E-mail: evelina9@mit.edu

1. Introduction

Knowledge of syntax enables people to string words together to convey an infinite range of new meanings. This productivity has a consequence, however: In constructing any given utterance, language users face numerous choices, including which words to use, how to order them, and how much detail to provide to communicate the intended meaning most efficiently. The factors that affect these production choices are numerous and include, most obviously, the meanings we want to convey, the conceptual accessibility of different entities (e.g., Bock, 1987; Ferreira, 1994; Gennari & MacDonald, 2009; MacWhinney, 1977; McDonald, Bock, & Kelly, 1993), the relationship between an utterance and its preceding context (e.g., Jackendoff, 1972; Paul, 1880), and the properties of our "audience" (e.g., Brennan, Galati, & Kuhlen, 2010; Jaeger, 2010). One other factor that has been hypothesized to shape our utterances, especially with respect to word order, is the amount of working memory resources required to plan an utterance (e.g., Gibson, 1998; Hawkins, 1994, 2004); for example, structures that contain nonlocal syntactic dependencies—where the element initiating a dependency has to be maintained in memory while intervening material is being produced—may be more costly for language producers. This hypothesis makes two predictions: (a) producers should prefer to use less complex structures; and (b) in cases where producers end up using the more complex structures, they should experience greater processing difficulty. Support for the first prediction comes from both corpus and experimental investigations. For example, Temperley (2007) found evidence of a preference for shorter dependencies across a wide range of constructions in the Penn Treebank corpus of English (see also Gildea & Temperley, 2010). In a picture-description task, Gennari, Mirkovic, and MacDonald (2012; see also Gennari & MacDonald, 2009) found that participants often choose syntactically simpler structures (e.g., a subject-extracted passive relative clause, which contains only local dependencies, instead of an object-extracted active relative clause, which contains a non-local dependency, as illustrated in (1) below). We here evaluate the second prediction, that of a cost associated with producing more complex structures, in an elicited production paradigm.

The constructions we investigate contain wh-dependencies (relative clauses in Experiment 1 and wh-questions in Experiment 2; e.g., Ross, 1967; Gibson, Piantadosi, Ichinco, & Fedorenko, 2012). These constructions have been extensively studied in the domain of language comprehension (see, e.g., O'Grady, 2011; Gibson, Tily, & Fedorenko, 2013, for recent overviews). From a theoretical standpoint, they represent one of the more complex syntactic phenomena. And from a practical standpoint, they allow for the investigation of syntactic complexity effects while controlling for lexical and plausibility factors. For example, the difference between the subject-extracted relative clause in (1a) and the object-extracted relative clause in (1b) is that the latter involves a non-local dependency between "attacked" and its object "who" (co-indexed with "the reporter"), while the former involves only local dependencies. The words are identical across the two sentences, and plausibility can be matched, either by keeping the meaning constant across constructions (e.g., by using "The senator who the reporter attacked" in the object-extracted condition) or by including four versions of each item, rotating the two noun phrases across syntactic positions.

(1) a. Subject-extracted RC

The reporter [who __ attacked the senator] admitted the error.

b. Object-extracted RC

The reporter [who the senator attacked __] admitted the error.

The syntactically more complex object-extractions like (1b) are consistently found to incur a greater cost in comprehension compared to their subject-extracted counterparts. Two main classes of accounts have been proposed in the sentence comprehension literature to explain the difficulty associated with object-extractions. According to *memory-based* accounts, structures with non-local dependencies require more working memory resources because the first element of the dependency needs to be retrieved from memory when the second element is encountered (e.g., Gibson, 1998, 2000; Grodner & Gibson, 2005; King & Just, 1991; Wanner & Maratsos, 1978), and this retrieval operation may further be susceptible to interference from intervening distractors (e.g., Gordon et al., 2001; Lewis et al., 2006; McElree et al., 2003). In contrast, *experience-based* accounts attribute difficulty with such structures to their relative rarity in the input (e.g., Gennari & MacDonald, 2008, 2009; Hale, 2001; Levy, 2008; Roland, Dick, & Elman, 2007; Wells et al., 2009).

The available comprehension data cannot be fully explained by either class of accounts alone. For example, memory-based accounts successfully predict the locus of comprehension difficulty in object-extractions (on the verb within the relative clause), because this is where the retrieval operation takes place (see, e.g., Grodner & Gibson, 2005; for discussion). Experience-based accounts predict difficulty as soon as the comprehender knows he or she has encountered an object-extraction, but little or no difficulty is observed at the subject noun phrase (cf. Staub, 2010). In contrast, experience-based, but not memory-based, accounts straightforwardly explain differences in processing difficulty depending on the types of noun phrases involved. For example, Reali and Christiansen (2007) have shown that if the relative-clause-internal noun phrase is replaced with a pronoun (e.g., "The senator that attacked you..."/"The senator that you attacked..."), the comprehension difficulty pattern reverses, such that the subject-extracted condition is now more costly (see also Troyer, O'Donnell, Fedorenko, & Gibson, 2011; for evidence from a production priming study). This effect is in line with the distribution of these types of relative clauses in corpora (e.g., Roland et al., 2007). Similarly, Traxler, Morris, and Seely (2002) have shown that the difficulty associated with object-extracted relative clauses is greatly reduced when the most frequent animacy configuration is used (i.e., animate agent and inanimate patient; e.g., "The movie that the producer saw..."). Given the complexity of the empirical picture, most researchers currently maintain that both a memory component and a probabilistic grammar component are needed in any complete model of language understanding (e.g., Demberg & Keller, 2008; Fedorenko, Woodbury, & Gibson, 2013; Gennari & MacDonald, 2009; Gibson et al., 2013; Levy, Fedorenko, & Gibson, 2013; Staub, 2010; Vasishth & Drenhaus, 2011).

In summary, our experience with language—from the level of simple n-grams all the way to complex long-distance dependencies—undoubtedly affects our ability to acquire and process different constructions (e.g., Fitz, Chang, & Christiansen, 2011). Object-extracted relative clauses with full noun phrases are infrequent in natural language, and their rarity plausibly explains some aspects of their comprehension difficulty. However, this rarity requires an explanation itself, and production difficulty may provide one such explanation (e.g., Bock & Levelt, 1994; Ferreira, 1996; Ferreira & Dell, 2000; Gibson & Pearlmutter, 1994). As discussed above, structures with non-local dependencies like the object-extraction in (1b) may be more costly to produce because one of the dependencies initiated by the pronoun "who" needs to be maintained in memory while the intervening material ("the senator" in [1b]) is being uttered. In the current experiments we seek to test whether syntactically complex object-extracted structures indeed cause more difficulty in language production than syntactically simpler subject-extracted structures.

2. Experiment 1: Relative clauses

2.1. Method

2.1.1. Participants

Twenty-four MIT undergraduates and members of the surrounding community participated for payment. All were native English speakers. All participants signed an informed consent form, in accordance with the regulations of the Internal Review Board at MIT.

2.1.2. Design, materials, and procedure

Forty items were created, with two versions of each. (A full list of experimental items appears in Appendix A.) Items were selected such that the intended production featured a relative clause with two definite, singular, animate noun phrases (NPs) (e.g., *the florist* and *the caterer*) involved in a reversible action (indicated by the verb; e.g., *mock*); no nouns or verbs were repeated across items. We chose to use full animate noun phrases (cf. a mixture of animate and inanimate nouns, or pronouns) because relative clauses with such NPs (a) exhibit a frequency difference in corpora, with subject-extracted relative clauses being much more frequent (e.g., Fitz et al., 2011; Roland et al., 2007), and (b) consistently yield comprehension complexity effects for object-extracted constructions.

An item consisted of a context sentence, three nouns denoting three entities, and a question targeting one of the nouns. Each item was presented in PowerPoint through a series of five screens (Fig. 1). The first screen contained a fixation cross (+), and participants were instructed to press the spacebar when they were ready to begin. After the spacebar press, the second screen appeared containing a context sentence. The context sentence introduced three individuals, always including two individuals of the same kind

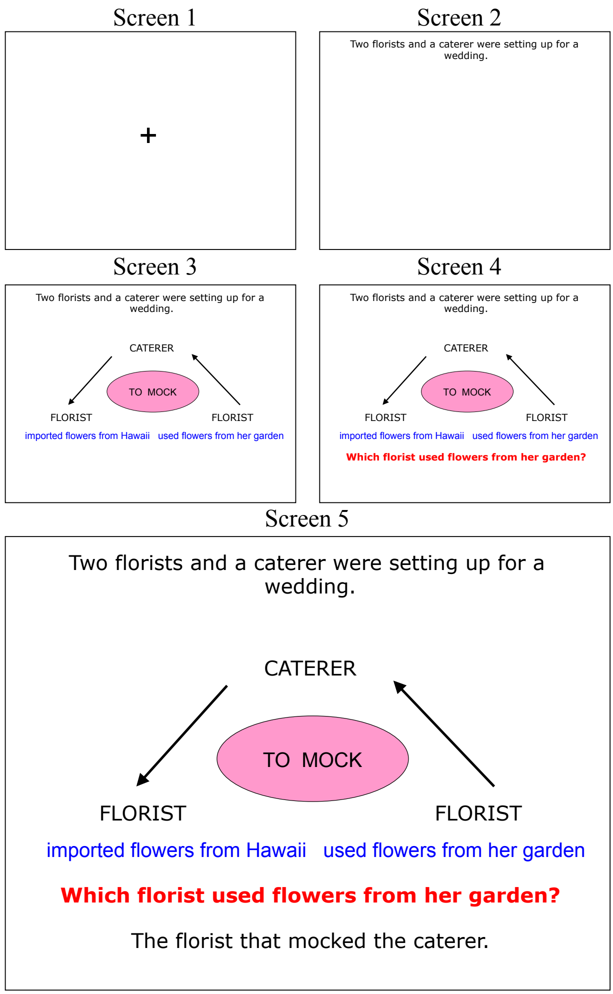

Fig. 1. Sample subject-extracted relative clause item from Experiment 1. In the object-extracted version of this item the question was "Which florist imported flowers from Hawaii?"

and one individual of a different kind (e.g., *Two florists and a caterer were setting up for a wedding*). After 4 seconds, the third screen displayed the nouns corresponding to the three entities introduced in the context sentence, the relevant verb, two arrows marking the relationships among the entities with respect to the verb, and two predicates distinguishing the two entities of the same kind (*the florists* in Fig. 1). The entities always appeared in a triangular configuration with the two nouns of the same kind at the bottom of the triangle and one of a different kind (e.g., *the caterer* in Fig. 1) at the top. The verb was presented in the infinitival form in a pink oval in the middle of the triangular configuration. Arrows indicated the agents and the patients in the scenario: the entity at the top of the configuration was always an agent of the verb with respect to one of the entities at the bottom, and a patient with respect to the other. The roles of the nouns at the bottom of the configuration were balanced across the trials within a list (half of the items had the agent on the left and half had the agent on the right), and within items across lists (such that across participants, both versions of a given item were presented). After another 4 seconds, the fourth screen displayed a question in red at the bottom of the screen, accompanied by a beep. The question asked about one of the nouns at the bottom of the triangular configuration (e.g., *Which florist used flowers from her garden?* in Fig. 1). Participants were instructed to answer the question with respect to the verb in the oval. Critically, the individual that was asked about was either the agent or the patient of the verb. When it was the agent of the verb (as in Fig. 1), the target production was a noun phrase modified by a subject-extracted relative clause (e.g., *the florist that mocked the caterer*). When the individual that was asked about was the patient of the verb, the target production was a noun phrase modified by an object-extracted relative clause (e.g., *the florist that the caterer mocked*). For example, in the object-extracted version of the item shown in Fig. 1 the targeted question would ask *Which florist imported flowers from Hawaii?*

Participants were told that they should begin speaking as soon as they had formulated their response, once they heard the beep accompanying the fourth screen. Once the response was produced, participants were told to press the spacebar again. The correct response was then displayed below the question on the fifth screen.

Participants were asked to avoid use of the passive voice (e.g., *the florist that was mocked by the caterer*). This instruction was given because previous sentence completion studies found that a passive construction is a frequent choice in situations where an object-extracted relative clause could be produced (e.g., Gennari & MacDonald, 2009), consistent with the general tendency to minimize dependency lengths in production (e.g., Temperley, 2007). After participants were familiarized with the instructions, they completed four practice trials.

Experimental trials were distributed across two lists, so that any given list contained only one version of each item. Each list contained 20 subject-extracted and 20 object-extracted trials. Four random orders of the trials in each list were created and the same number of participants saw each list and order.

2.1.3. Analyses

Correct productions corresponded to one of the two frames below, depending on whether they belonged to the subject- or object-extracted condition.

$$\textit{Subject-extracted}: \text{(latency) } \textit{the}_1\ \textit{NP}_1 \text{ that V-ed } \textit{the}_2\ \textit{NP}_2$$

$$\textit{Object-extracted}: \text{(latency) } \textit{the}_1\ \textit{NP}_1 \text{ that } \textit{the}_2\ \textit{NP}_2 \text{ V-ed}$$

Based on these frames, we defined seven regions of analysis within each production: *latency* (time from the onset of the beep to the onset of speech), the_1, NP_1, *that*, *V*, the_2, NP_2.

The productions were independently transcribed by two research assistants, and a third research assistant checked the work of the first two and made final decisions regarding points of disagreement. Correct responses were those that contained an uninterrupted production of the intended frame; disfluencies, incorrect words, and repetitions of words within the frame yielded an incorrect response. Substitution of the complementizer *that* with *who* was not counted as incorrect. (There were 39 instances of *who* in the subject-extracted condition, and 18 instances in the object-extracted condition, consistent with the higher frequency of *who* in subject-extracted structures in corpora; e.g., Roland et al., 2007.)

Of the 950 relative clause utterances produced, 823 (87%) were correct responses. Incorrect responses were relatively evenly distributed across the two conditions: 52% for subject-extracted targets and 48% for object-extracted targets. False start utterances, which began incorrectly but ended with an uninterrupted production of the intended frame, were included in most of the analyses below, and accounted for 3% of the correct responses. Of these false starts, 28% occurred in the subject-extracted condition and 72% in the object-extracted condition (mixed logit model predicting false start by condition, with random intercepts for subject and item and random slopes for condition grouped by subject and item: $\beta = -1.49$, $z = -2.83$, $p < .01$). Eight of the 18 object-extracted false starts (44%) began with a passive construction. (Utterances with false starts were not included in the analyses of total utterance duration.)

The boundaries of the seven regions were extracted by forcing alignments between the audio files and their associated transcripts (Shen, Strassel, & Cieri, 2007; Shen, White, & Hazen, 2009). These boundaries were then hand-checked and adjusted where needed by one of the authors (EL). Because only a handful of previous studies have examined syntactic complexity effects in production (e.g., Gennari & MacDonald, 2009; Gennari et al., 2012; Slevc, 2011), with none, to the best of our knowledge, exploring the acoustic properties of the productions, we chose a set of acoustic measures that have been previously shown to vary depending on the syntactic and semantic properties of sentence elements (see, e.g., Breen, Fedorenko, Wagner, & Gibson, 2010; for a recent overview). In particular, we extracted four acoustic measures from each of the defined regions using the Praat program (Boersma & Weenink, 2006): *duration*, *mean F0* (pitch), *maximum F0*, and *maximum intensity*. The measures were compared for each region across the two conditions using a linear mixed effects model predicting the relevant measure by condition (*subject-extracted* vs. *object-extracted*); because the *latency* region consisted of silence, only the *duration* measure was analyzed for this region. The models included random intercepts and slopes for subjects and items grouped by condition. Significance (p) values were estimated by using χ^2 likelihood tests to compare each model with a model identical in all respects except for the fixed effect of interest (i.e., condition), as recommended in Barr, Levy, Scheepers, and Tily (2013).

2.2. Results

Total utterance duration, estimated by summing over the durations of the seven regions defined above, differed significantly across conditions ($\beta = -0.098$, $t = -3.73$, $p < .001$); subject-extracted structures were produced with reliably shorter durations than their object-extracted counterparts (5.33 s vs. 5.94 s). (Duration measures were log transformed for all of the statistical tests reported.) Next, we report the results of statistical tests performed on regions internal to the utterances.

Only the first four regions (*latency,* the_1, NP_1, *that*) share the same utterance position across conditions; the other regions vary depending on condition with respect to their position within the utterance. For example, the *V* region in a subject-extracted production occurred fifth in the utterance, whereas in an object-extracted production it occurred seventh (utterance-finally). To avoid possible positional effects, we report only the statistical results for the first four regions; for a full list of statistical comparisons, see Appendix C.

Duration of the *latency* region differed significantly across conditions ($\beta = -0.117$, $t = -3.98$, $p < .001$); *latency* for object-extracted relative clauses was significantly longer than for subject-extracted relative clauses (3.53 s vs. 3.06 s). (We included false start utterances in this comparison. For those cases, we examined the time from the onset of the beep to the onset of the false start. The significance of the effect is not affected by the exclusion of these trials.) The same pattern of duration effects was observed for the_1 ($\beta = -0.089$, $t = -4.16$, $p < .001$) and for NP_1 ($\beta = -0.066$, $t = -3.90$, $p < .001$); both regions were produced faster in the subject-extracted condition than in the object-extracted condition (120 ms vs. 130 ms for the_1; 530 ms vs. 550 ms for NP_1). There was no significant durational difference in the production of *that* across conditions ($\beta = -0.044$, $t = -0.98$, $p = .326$). For the remaining acoustic measures—*mean F0*, *maximum F0*, *maximum intensity*—the only significant difference between conditions appears in the maximum intensity of the *that* region ($\beta = 0.818$, $t = 2.52$, $p < .05$); the maximum intensity of *that* in subject-extracted relative clauses was reliably larger than in object-extracted relative clauses.

2.3. Discussion

Subject-extracted relative clauses were produced with shorter latencies and durations and fewer disfluencies than object-extracted relative clauses, suggesting that producing object-extractions incurs a greater cognitive cost. Part of this cost might be explained by the instruction to avoid the passive construction; as mentioned above, passives are a common choice in production when an object-extracted relative clause could be produced (Gennari & MacDonald, 2009; Gennari et al., 2012; Montag & MacDonald, 2009). In our results, 44% of the false starts in object-extracted productions began with a passive and then switched to the intended, object-extracted relative clause construction. Consciously suppressing the passive construction in the object-extracted condition could have therefore led to the observed increase in latencies. However, the persistence of the durational differences for the sentence-medial words cannot be as straightforwardly explained by a passive-suppression strategy. Moreover, the avoidance of object-extractions and the attempts to produce passive structures suggest a cost of object-extraction independent of the message level, that is, at the level of structure. We return to this point in our discussion of Experiment 2.

No other acoustic measures showed reliable between-condition differences, except for a difference in maximum intensity on the complementizer region: *that* in subject-extracted relative clauses was produced with greater intensity than in object-extracted relative clauses. We leave the interpretation of this intensity effect to future work.

3. Experiment 2: Wh-questions

In Experiment 2, we sought to replicate the effects observed in Experiment 1 using another construction with wh-dependencies: wh-questions. In addition to generalizing the effects to a different construction, we wanted to make sure that our results in Experiment 1 were not due to some properties of the displays we used to elicit the productions. In particular, in the object-extracted condition of Experiment 1, participants may have had to shift their gaze in the direction opposite to the direction of the arrow to go from the target head noun to the target subject of the relative clause. The procedure for Experiment 2 was similar to that in Experiment 1, except that subject- and object-extracted wh-questions were elicited and a different visual display was used. If the durational differences and disfluency rates observed in Experiment 1 are due to the relative complexity of object-extractions, as opposed to some idiosyncratic property of relative clause structures or the visual display that elicited them, then we should find similar differences for wh-questions such that subject-extracted constructions are produced with shorter durations and fewer disfluencies than their object-extracted counterparts.

3.1. Method

3.1.1. Participants

The same 24 participants from Experiment 1 participated in Experiment 2. The order of the two experiments was counterbalanced across participants.

3.1.2. Design, materials, and procedure

Forty items were created, with two versions of each. (A full list of experimental items appears in Appendix B.) An item consisted of two nouns and a verb, together with a tense cue. Each item was presented in PowerPoint through a series of four screens (Fig. 2). The first screen contained a horizontal black bar, and participants were instructed to press the spacebar when they were ready to begin. After the spacebar press, a verb (*THANK* in Fig. 2) and a tense cue for the verb (*FUTURE* in Fig. 2) were displayed above the black bar. After 2 seconds, two nouns appeared below the black bar, accompanied by a beep. One of the nouns was preceded by the wh-determiner *which* (*Which LECTURER* in Fig. 2), whereas the other was bare (*GRADER* in Fig. 2). Furthermore, one of the nouns appeared in bold. Participants were told that the entity corresponding to the noun in bold was the agent of the action described by the verb (i.e., the thanker in Fig. 2), and the entity corresponding to the other (non-bolded) noun was the patient of the action (i.e., the thankee in Fig. 2). The two nouns were vertically arranged, and the

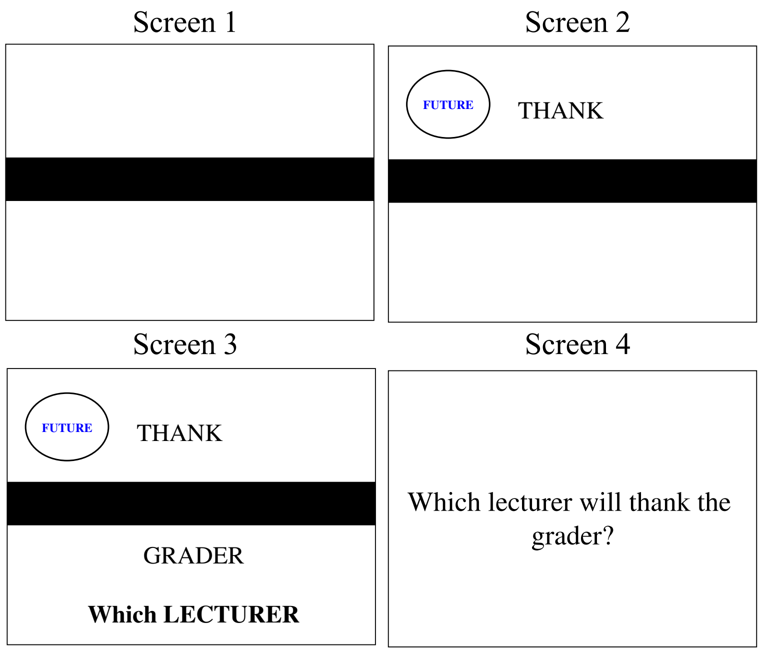

Fig. 2. Sample object-extracted WH item from Experiment 2. (The noun in bold is the agent, and the noun preceded by "which" has to initiate the question).

position (top vs. bottom) of the noun preceded by *which*, as well as of the noun in bold, was balanced across the trials within a list (a quarter of the items had the bold noun on top, preceded by *which*; another quarter had the bold noun on top and the non-bold noun preceded by *which*, etc.), and within items across lists (such that across participants all four versions of a given item were presented). Participants were instructed to formulate a question that would start with the *which*-phrase and that would ask about the relationship between the two nouns with respect to the verb. They were told to put the verb in the past or future tense, as indicated by the tense cue. Critically, the *which*-phrase was either the agent or the patient of the verb. When it was the agent of the verb, the target production was a subject-extracted wh-question (*Which lecturer will thank the grader?* in Fig. 2). And when the *which*-phrase was the patient of the verb, the target production was an object-extracted wh-question (*Which lecturer will the grader thank?*). Participants were told that the beep accompanying the third screen indicated that they could begin speaking whenever they were ready to produce the response. Once the response was produced, participants were told to press the spacebar again. The correct response then replaced the black bar. As in Experiment 1, participants were asked to avoid the use of the passive voice (e.g., *Which lecturer will be thanked by the grader?*). After participants were familiarized with the instructions, they completed eight practice trials.

Experimental trials were distributed across two lists, so that any given list contained only one version of each item. Each list contained 20 subject-extracted and 20 object-extracted trials. Four random orders of the trials in each list were created and the same number of participants saw each list and order.

3.1.3. Analyses

Correct productions corresponded to one of the four frames below, depending on whether they belonged to the subject- or object-extracted condition and whether they were produced in the past or future tense.

$$\textit{Subject-extracted PAST} : \text{(latency) which } NP_1 \text{ V-ed } \textit{the } NP_2$$

$$\textit{Object-extracted PAST} : \text{(latency) which } NP_1 \text{ did } \textit{the } NP_2\ V$$

$$\textit{Subject-extracted FUTURE} : \text{(latency) which } NP_1 \text{ will } V\ \textit{the } NP_2$$

$$\textit{Object-extracted FUTURE} : \text{(latency) which } NP_1 \text{ will } \textit{the } NP_2\ V$$

We defined seven regions within each production: *latency* (from the onset of the beep to the onset of speech), *which*, NP_1, *auxiliary* (corresponding to "did" or "will"), *V*, *the*, NP_2. The past tense version of the subject-extracted condition lacked an auxiliary, and so it contained only six regions for analysis. Across conditions, only the first three regions were identical in terms of their positions within the utterance.

Transcriptions and annotations were carried out in a manner identical to the first experiment. Of the 960 WH utterances produced, 865 were correct responses (90% correct). Incorrect responses were relatively evenly distributed across the two conditions: 41% subject-extracted, 59% object-extracted. False start utterances accounted for 9% of the correct responses. Of the false starts, 36% were subject-extracted and 64% were object-extracted (mixed logit model predicting false start by condition, with random intercepts for subject and item and random slopes for condition grouped by subject and item: $\beta = -0.64$, $z = -2.40$, $p < .05$). Three of the 49 object-extracted false starts (6%) began with a passive construction. Utterances with false starts were not included in the analysis of total utterance duration, which was estimated by summing over the durations of the regions internal to the utterance.

As in the previous experiment, region boundaries were determined by forcing alignments between the audio files and their transcripts, and they were subsequently hand-checked. For each region, we extracted four acoustic measures: *duration*, *mean F0*, *maximum F0*, and *maximum intensity*. Statistical comparisons were performed across conditions for each region using the same method described above for Experiment 1.

3.2. Results

Total utterance duration of future tense productions (which contained the same words across the subject- and object-extracted conditions) differed significantly across conditions ($\beta = -0.097$, $t = -2.90$, $p < .01$); subject-extractions were produced with reliably shorter durations than their object-extracted counterparts (4.77 s vs. 5.49 s). Next, we report the results of statistical tests performed on regions internal to the utterances.

Collapsing across the possible values for tense, only the first three regions (*latency, which,* NP_1) share the same utterance position across conditions. Within future tense utterances, the *auxiliary* region additionally overlaps in position. To avoid positional effects on the acoustic measures, we report only the statistical results for these regions; for a full list of statistical comparisons, consult Appendix C. As with relative clauses, duration of the *latency* region was significantly longer for object-extracted WH productions ($\beta = -0.201$, $t = -6.066$, $p < .001$) (2.23 s vs. 2.90 s). (As in Experiment 1, we included false start utterances in this comparison, but the significance of the effect is not affected by the exclusion of these trials.) Also, both the *which* region and the NP_1 region were produced with longer durations in the object-extracted condition ($\beta = -0.043$, $t = -2.52$, $p < .05$, *which*; $\beta = -0.082$, $t = -4.28$, $p < .001$, NP_1) (300 ms vs. 310 ms for *which*; 610 ms vs. 650 ms for NP_1). Within future tense productions, there was no durational difference on the *auxiliary* region ($\beta = 0.062$, $t = 1.45$, $p = .15$). Of the remaining acoustic measures, the only significant differences occur on the future tense *auxiliary* region: Both mean F0 and max intensity were greater for object-extracted productions ($\beta = -5.191$, $t = -3.21$, $p < .01$, *mean F0*; $\beta = -0.700$, $t = -2.91$, $p < .01$, *max intensity*).

3.3. Discussion

The results were similar to those in Experiment 1: Subject-extracted wh-questions are produced with shorter latencies and durations and fewer disfluencies than their object-extracted counterparts. It bears noting that whereas suppression of a passive construction may have contributed to the latency effect and the increased number of disfluencies for object-extracted relative clauses in Experiment 1, here this factor was less likely at play: Only 6% of the false starts in the object-extracted condition began with a passive construction (cf. 44% in Experiment 1). Furthermore, although it is difficult to rule out the potential contributions of gaze patterns (i.e., the need to make an eye movement in the direction opposite to the direction of the arrow in the object-extracted condition) to the observed effects in Experiment 1, in Experiment 2 a difference in eye movements is not likely given the similarity of the displays across conditions.

Additional acoustic differences occur on the *auxiliary* region, where a functional element occurs utterance medially (cf. the intensity effect on *that* from Experiment 1). We thus conclude that the increased production cost of object-extractions compared to subject-extractions characterizes the relevant syntactic property of both relative clauses and wh-questions.

4. Summary and conclusions

To summarize, across two experiments we observed that object-extracted structures (relative clauses in Experiment 1 and wh-questions in Experiment 2) caused greater difficulty in language production compared to subject-extracted structures. In particular, it took participants longer to initiate their productions and to articulate the words in the object-extracted conditions. No other acoustic measures (mean F0, maximum F0, and maximum intensity) produced consistent and interpretable patterns of results. Furthermore, participants were more likely to be disfluent in object-extracted conditions, as measured by the numbers of false starts. We therefore conclude that (a) syntactically complex object-extracted structures cause difficulty in language production relative to their subject-extracted counterparts and (b) longer durations and a greater number of disfluencies seem to be the most robust correlates of production difficulty, at least for syntactically complex structures.

As discussed in the introduction, experience-based theories and working-memory-based theories provide different, though not necessarily mutually exclusive, accounts of syntactic complexity effects in language comprehension. Experience-based accounts of language comprehension (Gennari & MacDonald, 2009; Hale, 2001; Levy, 2008) hold that structures like object-extracted relative clauses or wh-questions with full NPs may be more costly to understand because comprehenders do not encounter these structures as frequently as they do the corresponding subject-extracted structures (e.g., Duffield & Michaelis, 2011; Reali & Christiansen, 2007; Roland et al., 2007). However, such accounts raise the natural question of why such asymmetries exist in the first place. For example, consider the case of lexical frequencies: Lower frequency words take longer to process than higher frequency words, and presumably these effects are due to our differential experiences with higher versus lower frequency words. But what causes the distribution of lexical frequencies in the first place? Presumably, a major factor is that some meanings (e.g., "mother," "eat," "big") need to be communicated more frequently than other meanings (e.g., "cornucopia," "oscillate," "picayune"), consistent with similar frequency-meaning relationships—at least for higher-frequency words—cross-linguistically (Calude & Pagel, 2011).

The current results provide one possible explanation for the lower frequency of object-extracted constructions compared to subject-extracted constructions based on the relative complexity of producing these structures: Language producers may be avoiding constructions that are more costly to produce, opting for syntactically simpler constructions when possible (e.g., producing a passive subject-extracted relative clause instead of an active object-extracted relative clause). Still, the nature of this additional cost remains an open question. One possible source of difficulty may include active maintenance of the elements initiating syntactic dependencies (cf. the notion of "storage" in Gibson, 1998). Some evidence for the producers' difficulty of maintaining incomplete dependencies comes from a recent sentence completion study (Gibson & Fedorenko, 2011) where participants were shown sentence preambles initiating single or double center-embedded relative/complement clauses (e.g., *The professor who the student who the fact that ...*) and asked to continue them to make complete sentences. Although participants could correctly complete single center-embeddings (e.g., *The professor who the student...*), they often completed the double center-embedded conditions with just two of the three verbs, typically omitting the verb that would be associated with the middle noun. This result (a) is consistent with previous evidence from comprehension where participants have been shown to judge ungrammatical double center-embedded sentences (with two of the three required verbs) as acceptable as their grammatical counterparts with all three verbs, or even more acceptable (e.g., Gibson & Thomas, 1999), and (b) suggests that producers indeed have difficulty maintaining incomplete dependencies in planning utterances.

Another possible source of difficulty in producing object-extractions may have to do with susceptibility to similarity-based interference from noun phrases that intervene between the elements of a wh-dependency within the cue-based retrieval framework (Badecker & Lewis, unpublished data; Lewis & Vasishth, 2005; Van Dyke & Lewis, 2003). In the current studies we examined relative clauses and wh-questions with two animate full NPs, which share many features. Future work should consider a broader range of relative clause/wh-question configurations featuring various other noun types to see whether similar production difficulty is observed for object-extractions where the two NPs are less similar to each other.

The results presented here also add to ongoing discussions regarding the scope of incrementality in utterance planning and production (Allum & Wheeldon, 2007; Breen et al., 2010; Brown-Schmidt & Konopka, 2008; Brown-Schmidt & Tanenhaus, 2006; Ferreira & Swets, 2002; Garrett, 1980, 1982; Griffin & Bock, 2000; Martin, Crowther, Knight, Tamborello, & Yang, 2010; Schnur, 2011; Smith & Wheeldon, 1999). The fact that initial latencies and utterance-initial words are significantly slower in object-extractions relative to subject-extractions suggests that the planning of object-extracted dependencies and the cognitive costs associated with them also occur early, before speaking commences (see, e.g., Kuperman & Bresnan, 2012, for a recent discussion of acoustic variation and the time course of planning in production).

To conclude, syntactically complex structures, which contain nonlocal dependencies, are more difficult to plan and utter than syntactically simpler structures, which only contain local dependencies. It is therefore plausible that this difficulty contributes to the relative rarity of nonlocal dependencies (e.g., Collins, 1996; Temperley, 2007), which in turn explains some aspects of their comprehension difficulty.

Acknowledgments

We would like to thank several research assistants for help in setting up the experiments, and collecting and coding the data: Teresa Giblin, Diane Lo, Nicki Longe, and Alison Park. We thank Wade Shen (Lincoln Labs) for help with the forced alignment procedure, and Michael Wagner (McGill University) for providing the Praat script for extracting the acoustic features. We also thank Steve Piantadosi (University of Rochester) for help with the Praat script and Peter Graff (Intel Corporation) for help with the statistical analyses. We are grateful to the audience at CUNY 2007, TedLab members, Roger Levy, and Maria Polinsky for helpful comments. Finally, Ted Gibson deserves very special thanks for his help with all stages of the project and for his comments on the manuscript. The material in this paper is based in part on work done while the second author was serving as an NSF Program Director. Any opinions, findings, and conclusions expressed in this paper are those of the authors and do not necessarily reflect the views of the U.S. National Science Foundation. EF was supported by the Eunice Kennedy Shriver NICHD K99 award HD-057522.

References

Allum, P., & Wheeldon, L. (2007). Planning scope in spoken sentence production: The role of grammatical units. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 791–810.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278.

Bock, J. K. (1987). An effect of the accessibility of word forms on sentence structures. *Journal of Memory and Language*, *26*, 119–137.

Bock, J. K., & Levelt, W. J. M. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945–984). Orlando, FL: Academic Press.

Boersma, P., & Weenink, D. (2006). Praat: Doing phonetics by computer (Version 4.3.10) [Computer program]. Available at: http://www.praat.org/. Accessed October 5, 2008.

Breen, M., Fedorenko, E., Wagner, M., & Gibson, E. (2010). Acoustic correlates of information structure. *Language and Cognitive Processes*, *25*, 1044–1098.

Brennan, S., Galati, A., & Kuhlen, A. K. (2010). Two minds, one dialog: Coordinating speaking and understanding. In B. H. Ross (Ed.), *The psychology of learning and motivation*. Vol. *53* (pp. 301–344). Burlington, VT: Academic Press.

Brown-Schmidt, S., & Konopka, A. E. (2008). Little houses and casas pequeñas: Message formulation and syntactic form in unscripted speech with speakers of English and Spanish. *Cognition*, *109*, 274–280.

Brown-Schmidt, S., & Tanenhaus, M. K. (2006). Watching the eyes when talking about size: An investigation of message formulation and utterance planning. *Journal of Memory and Language*, *54*, 592–609.

Calude, A. S., & Pagel, M. (2011). How do we use language? Shared patterns in the frequency of word use across 17 word languages. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *366*, 1101–1107.

Collins, M. (1996). A new statistical parser based on bigram lexical dependencies. In *Proceedings of the 34th annual meeting of the ACL* (pp. 184–191). Stroudsburg, PA: Association for Computational Linguistics.

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

Duffield, C. J., & Michaelis, L. A. (2011). Why subject relatives prevail: Constraints versus constructional licensing. *Language and Cognition*, *3*, 171–208.

Fedorenko, E., Woodbury, R., & Gibson, E. (2013). Direct evidence of memory retrieval as a source of difficulty in long-distance structural dependencies in language. *Cognitive Science*, *37*, 378–394.

Ferreira, F. (1994). Choice of passive voice is affected by verb type and animacy. *Journal of Memory and Language*, *33*, 715–736.

Ferreira, V. S. (1996). Is it better to give than to donate? Syntactic flexibility in language production. *Journal of Memory and Language*, *35*, 724–755.

Ferreira, V. S., & Dell, G. S. (2000). The effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology*, *40*, 296–340.

Ferreira, F., & Swets, B. (2002). How incremental is language production? Evidence from the production of utterances requiring the computation of arithmetic sums. *Journal of Memory and Language*, *46*, 57–84.

Fitz, H., Chang, F., & Christiansen, M. H. (2011). A connectionist account of the acquisition and processing of relative clauses. In E. Kidd (Ed.), *The acquisition of relative clauses*. Vol. *8* (pp. 39–60). Amsterdam: John Benjamins.

Garrett, M. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language production*. Vol. *1* (pp. 177–220)., *Speech and talk* London: Academic Press.

Garrett, M. (1982). Production of speech: Observations from normal and pathological language use. In A. Ellis (Ed.), *Normality and pathology in cognitive functions* (pp. 19–76). New York: Academic Press.

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

Gennari, S., & MacDonald, M. (2009). Linking production and comprehension processes: The case of relative clauses. *Cognition*, *111*, 1–23.

Gennari, S., Mirkovic, J., & MacDonald, M. (2012). Animacy and competition in relative clause production: A cross-linguistic investigation. *Cognitive Psychology*, *65*, 141–176.

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* (pp. 95–126). Cambridge, MA: MIT Press.

Gibson, E., & Fedorenko, E. (2011). *The domain-generality of working memory resources for language*. Paris: AMLaP conference.

Gibson, E., & Pearlmutter, N. (1994). A corpus-based analysis of psycholinguistic constraints on PP attachment. In C. Clifton Jr, L. Frazier, & K. Rayner (Eds.), *Perspectives in sentence processing* (pp. 181–198). Hillsdale, NJ: Erlbaum.

Gibson, E., Piantadosi, S. T., Ichinco, D., & Fedorenko, E. (2012). *Evaluating structural overlap across constructions: Inter-subject analysis of co-variation*. Portland, OR: Talk presented at the Linguistic Society of America.

Gibson, E., & Thomas, J. (1999). Memory limitations and structural forgetting: The perception of complex ungrammatical sentences as grammatical. *Language and Cognitive Processes*, *14*, 225–248.

Gibson, E., Tily, H., & Fedorenko, E. (2013).The processing complexity of English relative clauses. In M. Sanz, I. Laka, & M. K. Tanenhaus (Eds.), *Language down the garden path: The cognitive and biological basis for linguistic structures*. Oxford, England: Oxford University Press.

Gildea, D., & Temperley, D. (2010). Do grammars minimize dependency length? *Cognitive Science*, *34*, 286–310.

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

Griffin, Z., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science*, *11*, 274–279.

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

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

Hawkins, J. (1994). *A performance theory of order and constituency*. Cambridge, UK: Cambridge University Press.

Hawkins, J. (2004). *Efficiency and complexity in grammars*. Oxford, England: Oxford University Press.

Jackendoff, R. (1972). *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.

Jaeger, T. F. (2010). Redundancy and reduction: Speakers manage information density. *Cognitive Psychology*, *61*, 23–62.

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

Kuperman, V., & Bresnan, J. (2012). The effects of construction probability on word durations during spontaneous incremental sentence production. *Journal of Memory and Language*, *66*, 588–611.

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

Levy, R., Fedorenko, E., & Gibson, E. (2013). The syntactic complexity of Russian relative clauses. *Journal of Memory and Language*, *69*, 461–495.

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

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

MacWhinney, B. (1977). Starting points. *Language*, *53*, 152–168.

Martin, R., Crowther, J., Knight, M., Tamborello, F., & Yang, C.-L. (2010). Planning in sentence production: Evidence for the phrase as a default planning scope. *Cognition*, *116*, 177–192.

McDonald, J. L., Bock, J. K., & Kelly, M. H. (1993). Word and world order: Semantic, phonological, and metrical determinants of serial position. *Cognitive Psychology*, *25*, 188–230.

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

Montag, J. L., & MacDonald, M. (2009). Word order doesn't matter: Relative clause production in English and Japanese. In N. A. Taatgen & van Rijn H. (Eds.), *Proceedings of the 31th Annual Conference of the Cognitive Science Society* (pp. 2594–2599). Hillsdale, NJ: Lawrence Erlbaum Associates.

O'Grady, W. (2011). Relative clauses: Processing and acquisition. In E. Kidd (Ed.), *The acquisition of relative clauses: Processing, typology and function* (pp. 13–38). Amsterdam: John Benjamins.

Paul, H. (1880). *Prinzipien der sprachgeschichte [Principles of the history of language]*. Tübingen, Germany: Max Niemeyer.

Reali, F., & Christiansen, M. (2007). Processing of relative clauses is made easier by frequency of occurrence. *Journal of Memory and Language*, *57*, 1–23.

Roland, D., Dick, F., & Elman, J. L. (2007). Frequency of basic English grammatical structures: A corpus analysis. *Journal of Memory and Language*, *57*, 348–379.

Ross, J. R. (1967). Constraints on variables in syntax. Doctoral dissertation, Massachusetts Institute of Technology.

Schnur, T. T. (2011). Phonological planning during sentence production: Beyond the verb. *Frontiers in Psychology*, *2*(319), 1–15.

Shen, W., Strassel, S., & Cieri, C. (2007). Software Tools for Transcription and Annotation of Interview Recordings. NWAV 36.

Shen, W., White, C. M., & Hazen, T. J. (2009). A comparison of query-by-example methods for spoken term detection. *INTERSPEECH 2009*, *4*, Red Hook, NY: Curran Associates, Inc. 2107–2110.

Slevc, L. R. (2011). Saying what's on your mind: Working memory effects on sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 1503–1514.

Smith, M., & Wheeldon, L. (1999). High level processing scope in spoken sentence production. *Cognition*, *73*, 205–246.

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. 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.

Troyer, M., O'Donnell, T. J., Fedorenko, E., & Gibson, E. (2011). Storage and computation in syntax: Evidence from relative clause priming. In N.A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 33rd annual meeting of the Cognitive Science Society* (pp. 336–341). Austin, TX: Cognitive Science Society.

Van Dyke, J., & Lewis, R. (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.

Vasishth, S., & Drenhaus, H. (2011). Locality effects in German. *Dialogue and Discourse*, *2*(1), 59–82.

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.
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.

Appendix A: Experiment 1 materials

Below we provide the target productions for the subject- and object-extracted versions of each of the 40 experimental items in Experiment 1.

Item	Cond.	Target Production
1	Object	The vocalist that the dancer complimented
	Subject	The vocalist that complimented the dancer
2	Object	The dealer that the curator interrogated
	Subject	The dealer that interrogated the curator
3	Object	The merchant that the farmer paid
	Subject	The merchant that paid the farmer
4	Object	The singer that the playwright called
	Subject	The singer that called the playwright
5	Object	The bartender that the gangster stabbed
	Subject	The bartender that stabbed the gangster
6	Object	The aristocrat that the king poisoned
	Subject	The aristocrat that poisoned the king
7	Object	The actor that the producer shoved
	Subject	The actor that shoved the producer
8	Object	The guard that the warden searched for
	Subject	The guard that searched for the warden
9	Object	The bridesmaid that the groom toasted
	Subject	The bridesmaid that toasted the groom
10	Object	The soprano that the pianist thanked
	Subject	The soprano that thanked the pianist
11	Object	The freshman that the senior nominated
	Subject	The freshman that nominated the senior
12	Object	The attorney that the judge questioned
	Subject	The attorney that questioned the judge
13	Object	The bellboy that the maid blamed
	Subject	The bellboy that blamed the maid
14	Object	The sister that the bride photographed
	Subject	The sister that photographed the bride
15	Object	The crook that the informant whispered to
	Subject	The crook that whispered to the informant
16	Object	The chick that the hen peeped at
	Subject	The chick that peeped at the hen
17	Object	The performer that the manager looked for
	Subject	The performer that looked for the manager

(continued)

Appendix A. (continued)

Item	Cond.	Target Production
18	Object	The girl that the boy slapped
	Subject	The girl that slapped the boy
19	Object	The politician that the talk-show host insulted
	Subject	The politician that insulted the talk-show host
20	Object	The chipmunk that the squirrel chased
	Subject	The chipmunk that chased the squirrel
21	Object	The burglar that the detective shot
	Subject	The burglar that shot the detective
22	Object	The boy that the father tackled
	Subject	The boy that tackled the father
23	Object	The cellist that the soloist yelled at
	Subject	The cellist that yelled at the soloist
24	Object	The candidate that the moderator berated
	Subject	The candidate that berated the moderator
25	Object	The novice that the grandmaster defeated
	Subject	The novice that defeated the grandmaster
26	Object	The model that the designer congratulated
	Subject	The model that congratulated the designer
27	Object	The therapist that the psychologist laughed at
	Subject	The therapist that laughed at the psychologist
28	Object	The baseball player that the coach pushed
	Subject	The baseball player that pushed the coach
29	Object	The student that the teacher complained to
	Subject	The student that complained to the teacher
30	Object	The football player that the cheerleader flattered
	Subject	The football player that flattered the cheerleader
31	Object	The employee that the boss interrupted
	Subject	The employee that interrupted the boss
32	Object	The analyst that the executive praised
	Subject	The analyst that praised the executive
33	Object	The brother that the toddler pinched
	Subject	The brother that pinched the toddler
34	Object	The customer that the barber entertained
	Subject	The customer that entertained the barber
35	Object	The American player that the German player fouled
	Subject	The American player that fouled the German player
36	Object	The reporter that the senator attacked
	Subject	The reporter that attacked the senator
37	Object	The musician that the conductor put down
	Subject	The musician that put down the conductor
38	Object	The cowboy that the rancher helped
	Subject	The cowboy that helped the rancher
39	Object	The sailor that the captain hugged
	Subject	The sailor that hugged the captain
40	Object	The busboy that the waitress assisted
	Subject	The busboy that assisted the waitress

Appendix B: Experiment 2 materials

Below we provide the target productions for the future tense subject- and object-extracted versions of each of the 40 experimental items in Experiment 2.

Item	Tense	Cond.	Target Production
1	Future	Object	Which student will the teacher ignore?
	Future	Object	Which teacher will the student ignore?
	Past	Subject	Which student ignored the teacher?
	Past	Subject	Which teacher ignored the student?
2	Future	Subject	Which con-man will accuse the robber?
	Future	Subject	Which robber will accuse the con-man?
	Past	Object	Which con-man did the robber accuse?
	Past	Object	Which robber did the con-man accuse?
3	Future	Object	Which bachelor will enrange the neighbor?
	Future	Object	Which neighbor will enrage the bachelor?
	Past	Subject	Which bachelor did the neighbor enrage?
	Past	Subject	Which neighbor did the bachelor enrage?
4	Future	Subject	Which accountant will the programmer e-mail?
	Future	Subject	Which programmer will the accountant e-mail?
	Past	Object	Which accountant e-mailed the programmer?
	Past	Object	Which programmer e-mailed the accountant?
5	Future	Object	Which governor will the politician entertain?
	Future	Object	Which politician will the governor entertain?
	Past	Subject	Which governor entertained the politician?
	Past	Subject	Which politician entertained the governor?
6	Future	Subject	Which colonel will hit the sniper?
	Future	Subject	Which sniper will hit the colonel?
	Past	Object	Which colonel did the sniper hit?
	Past	Object	Which sniper did the colonel hit?
7	Future	Subject	Which composer will envy the pianist?
	Future	Subject	Which pianist will envy the composer?
	Past	Object	Which composer did the painist envy
	Past	Object	Which pianist did the composer envy?
8	Future	Object	Which terrorist will the traitor frighten?
	Future	Object	Which traitor will the terrorist frighten?
	Past	Subject	Which terrorist frightened the traitor?
	Past	Subject	Which traitor frightened the terrorist?

(continued)

Appendix B. (continued)

Item	Tense	Cond.	Target Production
9	Future	Object	Which sergeant will the veteran respect?
	Future	Object	Which veteran will the sergeant respect?
	Past	Subject	Which sergeant respected the veteran?
	Past	Subject	Which veteran respected the sergeant?
10	Future	Subject	Which gentleman will avoid the lady?
	Future	Subject	Which lady will avoid the gentleman?
	Past	Object	Which gentleman did the lady avoid?
	Past	Object	Which lady did the gentleman avoid?
11	Future	Subject	Which knight will trust the lord?
	Future	Subject	Which lord will trust the knight?
	Past	Object	Which knight did the lord trust?
	Past	Object	Which lord did the knight trust?
12	Future	Object	Which congressman will the Democrat support?
	Future	Object	Which Democrat will the congressman support?
	Past	Subject	Which congressman supported the Democrat?
	Past	Subject	Which Democrat supported the congressman?
13	Future	Object	Which visitor will the woman hug?
	Future	Object	Which woman will the visitor hug?
	Past	Subject	Which visitor hugged the woman?
	Past	Subject	Which woman hugged the visitor?
14	Future	Subject	Which fisherman will rescue the swimmer?
	Future	Subject	Which swimmer will rescue the fisherman?
	Past	Object	Which fisherman did the swimmer rescue?
	Past	Object	Which swimmer did the fisherman rescue?
15	Future	Subject	Which ambassador will identify the translator?
	Future	Subject	Which translator will identify the ambassador?
	Past	Object	Which ambassador did the translator identify?
	Past	Object	Which translator did the ambassador identify?
16	Future	Object	Which dancer will the singer welcome?
	Future	Object	Which singer will the dancer welcome?
	Past	Subject	Which dancer welcomed the singer?
	Past	Subject	Which singer welcomed the dancer?
17	Future	Object	Which prowler will the sheriff see?
	Future	Object	Which sheriff will the prowler see?
	Past	Subject	Which prowler saw the sheriff?
	Past	Subject	Which sheriff saw the prowler?
18	Future	Subject	Which employee will phone the receptionist?
	Future	Subject	Which receptionist will phone the employee?
	Past	Object	Which employee did the receptionist phone?
	Past	Object	Which receptionist did the employee phone?
19	Future	Subject	Which customer will offend the saleslady?
	Future	Subject	Which saleslady will offend the customer?
	Past	Object	Which customer did the saleslady offend?
	Past	Object	Which saleslady did the customer offend?

(continued)

Appendix B. (continued)

Item	Tense	Cond.	Target Production
20	Future	Object	Which actor will the producer praise?
	Future	Object	Which producer will the actor praise?
	Past	Subject	Which actor praised the producer?
	Past	Subject	Which producer praised the actor?
21	Future	Object	Which boy will the girl date?
	Future	Object	Which girl will the boy date?
	Past	Subject	Which boy dated the girl?
	Past	Subject	Which girl dated the boy?
22	Future	Subject	Which criminal will kill the policeman?
	Future	Subject	Which policeman will kill the criminal?
	Past	Object	Which criminal did the policeman kill?
	Past	Object	Which policeman did the criminal kill?
23	Future	Subject	Which child will meet the tutor?
	Future	Subject	Which tutor will meet the child?
	Past	Object	Which child did the tutor meet?
	Past	Object	Which tutor did the child meet?
24	Future	Object	Which admiral will the soldier like?
	Future	Object	Which soldier will the admiral like?
	Past	Subject	Which admiral liked the soldier?
	Past	Subject	Which soldier liked the admiral?
25	Future	Object	Which lawyer will the mayor advise?
	Future	Object	Which mayor will the lawyer advise?
	Past	Subject	Which lawyer advised the mayor?
	Past	Subject	Which mayor advised the lawyer?
26	Future	Subject	Which professor will address the freshman?
	Future	Subject	Which freshman will address the profesor?
	Past	Object	Which freshman did the professor address?
	Past	Object	Which professor did the freshman address?
27	Future	Subject	Which grader will thank the lecturer?
	Future	Subject	Which lecturer will thank the grader?
	Past	Object	Which grader did the lecturer thank?
	Past	Object	Which lecturer did the grader thank?
28	Future	Object	Which psychologist will the surgeon consult?
	Future	Object	Which surgeon will the psychologist consult?
	Past	Subject	Which psychologist consulted the surgeon?
	Past	Subject	Which surgeon consulted the psychologist?
29	Future	Object	Which author will the publisher select?
	Future	Object	Which publisher will the author select?
	Past	Subject	Which author selected the publisher?
	Past	Subject	Which publisher selected the author?
30	Future	Subject	Which drummer will recommend the songwriter?
	Future	Subject	Which songwriter will recommend the drummer?
	Past	Object	Which drummer did the songwriter recommend?
	Past	Object	Which songwriter did the drummer recommend?

(continued)

Appendix B. (continued)

Item	Tense	Cond.	Target Production
31	Future	Subject	Which electrician will hire the mechanic?
	Future	Subject	Which mechanic will hire the electrician?
	Past	Object	Which electrician did the mechanic hire?
	Past	Object	Which mechanic did the electrician hire?
32	Future	Object	Which auditor will the bookkeeper confuse?
	Future	Object	Which bookkeeper will the auditor confuse?
	Past	Subject	Which auditor confused the bookkeeper?
	Past	Subject	Which bookkeeper confused the auditor?
33	Future	Object	Which general will the marine capture?
	Future	Object	Which marine will the general capture?
	Past	Subject	Which general captured the marine?
	Past	Subject	Which marine captured the general?
34	Future	Subject	Which comedian will hate the photographer?
	Future	Subject	Which photographer will hate the comedian?
	Past	Object	Which comedian did the photographer hate?
	Past	Object	Which photographer did the comedian hate?
35	Future	Subject	Which hostess will irritate the waiter?
	Future	Subject	Which waiter will irritate the hostess?
	Past	Object	Which hostess did the waiter irritate?
	Past	Object	Which waiter did the hostess irritate?
36	Future	Object	Which reporter will the senator intimidate?
	Future	Object	Which senator will the reporter intimidate?
	Past	Subject	Which reporter intimidated the senator?
	Past	Subject	Which senator intimidated the reporter?
37	Future	Object	Which banker will the investor bore?
	Future	Object	Which investor will the banker bore?
	Past	Subject	Which banker bored the investor?
	Past	Subject	Which investor bored the banker?
38	Future	Subject	Which mathematician will teach the physicist?
	Future	Subject	Which physicist will teach the mathematician?
	Past	Object	Which mathematician did the physicist teach?
	Past	Object	Which physicist did the mathematician teach?
39	Future	Subject	Which businessman will greet the inventor?
	Future	Subject	Which inventor will greet the businessman?
	Past	Object	Which businessman did the inventor greet?
	Past	Object	Which inventor did the businessman greet?
40	Future	Object	Which orderly will the resident observe?
	Future	Object	Which resident will the orderly observe?
	Past	Subject	Which orderly observed the resident?
	Past	Subject	Which resident observed the orderly?

Appendix C: Statistical analyses

Region	β	*SE*	*t*	χ^2	*p*
Log Duration					
RC					
lat	−0.12	0.03	−3.98	12.2	<.001
the_1	−0.09	0.02	−4.16	13.33	<.001
NP_1	−0.07	0.02	−3.9	12.74	<.001
that	−0.04	0.04	−0.98	0.97	.33
V	−0.22	0.03	−7.16	29.41	<.001
the_2	−0.01	0.02	−0.52	0.25	.62
NP_2	0.18	0.03	6.07	27.32	<.001
WH					
lat	−0.2	0.03	−6.07	23.5	<.001
wh	−0.04	0.02	−2.52	5.68	<.05
NP_1	−0.08	0.02	−4.28	13.8	<.001
aux ('will')	0.06	0.04	1.45	2.09	.15
V	−0.17	0.04	−4.56	17.56	<.001
the_2	−0.08	0.03	−2.42	5.35	<.05
NP_2	0.15	0.02	6.71	26.71	<.001
Mean Pitch (F0)					
RC					
the_1	−0.25	2.01	−0.12	0.02	.90
NP_1	−0.54	1.02	−0.53	0.28	.60
that	1.78	3.07	0.58	0.33	.57
V	10.7	4.87	2.2	4.56	<.05
the_2	−10.38	2.09	−4.97	17.47	<.001
NP_2	−11.86	3.71	−3.2	8.88	<.01
WH					
wh	−0.128	2.191	−0.058	0	1
NP_1	−0.41	0.97	−0.43	0.17	.68
aux ('will')	−5.19	1.62	−3.21	10.17	<.01
V	11.4	3.86	2.95	7.62	<.01
the_2	−6.88	2.97	−2.32	5.76	<.05
NP_2	−6.47	3.26	−1.99	3.82	.05
Max Pitch (F0)					
RC					
the_1	−1.07	3.42	−0.31	0.08	.78
NP_1	−0.19	1.92	−0.1	0.28	.60
that	3.03	4.46	0.68	0.46	.50
V	−3.95	5.69	−0.7	0.48	.49
the_2	−9.89	2.27	−4.36	16.32	<.001
NP_2	4.87	5.07	0.96	0.92	.34
WH					
wh	1.916	3.544	0.541	0.2853	.59
NP_1	−1.07	2.44	−0.44	0.19	.66
aux ('will')	−4.38	2.87	−1.53	2.33	.13

(continued)

Appendix C. (*continued*)

Region	β	*SE*	*t*	χ^2	*p*
V	1.55	4.61	0.34	0.1	.75
the_2	−9.02	3.8	−2.38	5.29	<.05
NP_2	5.77	4.02	1.44	2.04	.15
Max Intensity					
RC					
the_1	−0.42	0.22	−1.89	0.47	.06
NP_1	−0.01	0.19	−0.05	0	1
that	0.82	0.33	2.52	5.74	<.05
V	2.63	0.31	8.51	37.44	<.001
the_2	−1.56	0.33	−4.8	17.46	<.001
NP_2	−2.69	0.37	−7.18	31.39	<.001
WH					
wh	−0.3299	0.2163	−1.53	2.301	0.13
NP_1	−0.34	0.21	−1.65	2.64	0.10
aux ('will')	−0.7	0.24	−2.91	7.52	<.01
V	2.32	0.4	5.88	21.91	<.001
the_2	−2.47	0.3	−8.31	34.12	<.001
NP_2	−2	0.27	−7.34	28.86	<.001

($df = 1$ in all tests).