## Title

Storage and computation in syntax: Evidence from relative clause priming



**Journal**
Proceedings of the Annual Meeting of the Cognitive Science Society, 33(33)

**ISSN**
1069-7977

**Authors**
Troyer, Melissa
O'Donnell, Timothy J.
Fedorenko, Evelina
et al.

**Publication Date**
2011

Peer reviewed

# Storage and computation in syntax: Evidence from relative clause priming

**Melissa Troyer[1], Timothy J. O'Donnell[2], Evelina Fedorenko[1], and Edward Gibson[1]**
{mltroyer, evelina9, egibson} @mit.edu, Department of Brain and Cognitive Sciences, MIT[1]
timo@wjh.harvard.edu, Department of Psychology, Harvard University[2]

**Abstract**

In morphology, researchers have provided compelling evidence for the storage of even fully compositional structures that could otherwise be computed by rule. For example, a high-frequency word composed of multiple morphemes (e.g., root + plural inflection) may be stored directly rather than computed on the fly (e.g., Baayen, Dijkstra, & Schreuder, 1997). Here, we investigate similar questions of storage and computation in syntax, a domain in which evidence of storage of fully compositional structures has been less forthcoming. We approach this question using syntactic priming, a method exploiting the tendency of individuals to repeat syntactic structures that they have recently produced (Bock, 1986). As a test case, we investigate relative clauses (RCs). RCs are both abstract and syntactically complex but are nevertheless frequent in natural language (Roland, Dick, & Elman, 2007). Moreover, differences in processing complexity between different RC types are at least partially predicted by frequency (e.g., Reali & Christiansen, 2007). RCs are therefore an ideal domain to look for evidence of storage of abstract, compositional syntactic structure. If the structures underlying high-frequency RC types are stored and retrieved from memory as whole units instead of being computed online from smaller units, then these stored structures should be susceptible to priming. Across three experiments, we observed that priming of object-extracted RCs is sensitive to a) the type of noun phrase in the embedded subject position (a full NP vs. a pronoun), and b) the type of relative pronoun (*who* vs. *that*). These results suggest that the representations of some types of RCs involve storage of large units which include both syntactic and lexical information. We interpret these results as providing support for models of syntax that allow for complex mixtures of storage and computation.

**Keywords:** syntax, relative clauses, priming

## Introduction

An important open question in the study of language involves the nature of the syntactic representations that are stored in long-term memory. At one extreme, the language user might only store the smallest fragments of structure needed for composing meaning. Across different classes of grammatical theories these minimal fragments of structure have sometimes been represented with devices such as context-free grammar rules (e.g., S → NP VP; NP → Det N; Sag, Wasow, & Bender, 2003), basic combinatory types (Steedman, 2000), immediate word-word dependencies (e.g., Mel'čuk, 1988), or basic "merge" operations (e.g., Chomsky, 1995). Processing a sentence using only these minimal units would require accessing large numbers of these stored "items" (rules) from long-term memory and combining them on-the-fly to create or infer new complex meanings. At the other extreme, language users might store many or all of the utterances and combinations of structure that they have ever encountered, increasing the number of items that must be stored in memory but decreasing the amount of computation required to process utterances.

Where does natural language fall on this continuum? Structures that are idiosyncratic or non-compositional—such as idioms and monomorphemic words—must be stored (Bloomfield, 1933). A more controversial question is whether fully compositional structures are sometimes stored as well. In the domain of words, there is evidence for storage of even fully regular morphological structure across a variety of languages and morphological systems. Examples include English verbal morphology (Alegre & Gordon, 1999), English noun pluralization (Sereno & Jongman, 1997), Italian and Dutch noun pluralization (Baayen et al., 1997a; Baayen et al., 1997b) and Finnish case and number marking (Bertram et al., 1999).

In syntax, the evidence for storage of compositional structure is less certain. Although a number of recent theoretical proposals in linguistics and psycholinguistics have advocated the idea that fully compositional syntactic structures can be stored (e.g., Jackendoff, 2002; Goldberg, 2005) it has proven difficult to obtain experimental evidence for such storage. Several recent studies have shown evidence for storage of long sequences of words (Bod, 2001; Bannard and Matthews, 2008; Tremblay, 2009; Arnon & Snider, 2010). These results show that storage is pervasive, even for fully compositional structure above the word level. However, while these results are important and suggestive, they fall short of providing evidence for the storage of complex *abstract* syntactic structures.

Here, we provide experimental evidence that the linguistic system stores complex syntactic structures that combine both abstraction and lexical specificity. To do this we investigate object-extracted relative clause constructions (ORCs). The nature of ORCs, in particular, makes them an

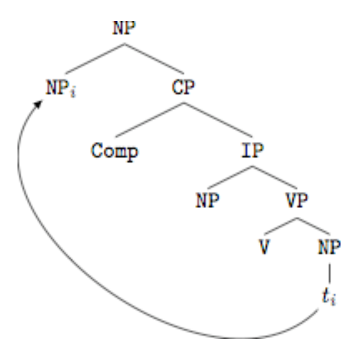


Figure 1: Abstract syntactic structure of an ORC

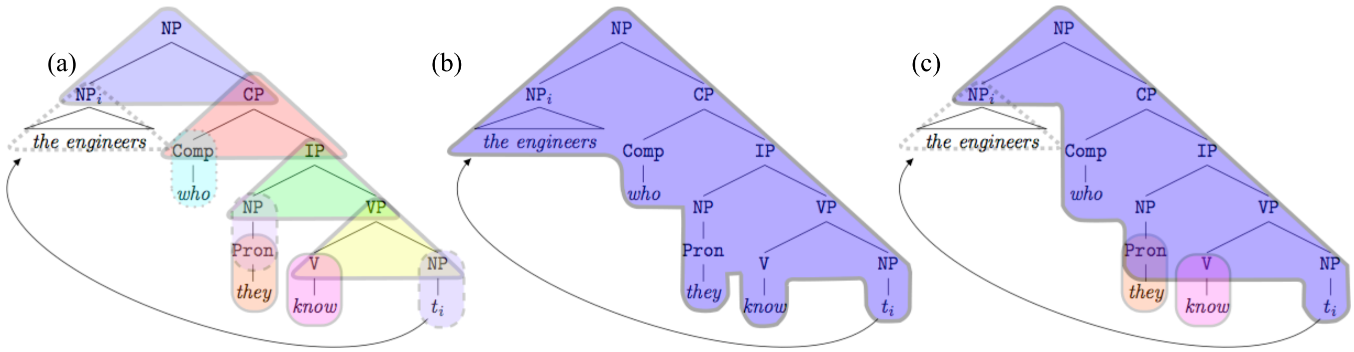


Figure 2: Granularity of potentially stored ORC structures

ideal test case for investigating the granularity of the syntactic units that are stored. An ORC modifies a head noun which has been extracted from the object position of the RC. For instance, in the sentence, "The man that I met yesterday was at the basketball game," the ORC "that I met yesterday" modifies the head noun "the man", which is co-referential with the empty object of the verb "met". RCs therefore involve an abstract structure with an embedded clause, a constituent that has been displaced, and several other syntactic positions (see Figure 1).

Previous work on relative clause attachment provides some evidence for priming (and storage) of abstract syntactic structure (Scheepers, 2003). However, one important question is whether the linguistic system stores information about specific lexical items together with this abstract structure. Fig. 2 illustrates the same ORC under three different storage strategies/hypotheses. These range from a strategy which is maximally combinatorial (Fig. 2a), to a strategy where all lexical information is stored together with the abstract syntactic structure (Fig. 2b), to an intermediate case including some lexical content but still allowing for variability in several of the syntactic positions (2c).

Our experiments contrasted two structural variants of ORCs: (1) ORCs with a pronoun versus a definite NP in the embedded subject position (see Figure 3a vs 3c), and (2) ORCs with *that* versus *who* as relative pronoun (see Figure 3a vs 3b).

Reali & Christiansen (2007) observed that ORCs with a personal pronoun in subject position (henceforth, "pronominal ORCs," e.g., Fig. 3a-3b) are more frequent than ORCs with a definite NP in subject position ("definite-NP ORCs," e.g. Fig. 3c). Reali and Christiansen also found that pronominal ORCs are read faster than matched pronominal SRCs, which are typically considered to be less syntactically complex (see also Warren & Gibson, 2002). We therefore hypothesized that greater priming of ORCs would be observed when the prime contains a pronominal ORC (compared to a definite-NP ORC).

Roland et al. (2007) further observed that ORCs occur more frequently following inanimate head NPs, compared to animate head NPs (see Traxler et al., 2002; Mak et al., 2002, for reading time evidence suggesting that ORCs with inanimate head NPs are processed faster than those with animate head NPs). Because the relative pronoun *that* can refer to either animate or inanimate NPs, but the relative pronoun *who* can only refer to animate NPs, we hypothesized that ORCs might be more likely to occur with the relative pronoun *that,* compared to *who.* This hypothesis was confirmed by a corpus study. We searched the Switchboard, Brown, and WSJ corpora from the Treebank-2 distribution (Marcus, Santorini, & Marcinkiewicz, 1995) and found that in each corpus, less than 1.5% of all ORCs occurred with *who*. In contrast, between 36% and 88% of ORCs in each corpus occurred with *that*. We therefore hypothesized that greater priming of ORCs would be observed when the prime contains an ORC with the relative pronoun *that* (compared to the relative pronoun *who*).

We conducted three sentence-completion experiments in which participants were instructed to form complete English sentences from short preambles (see e.g., Scheepers, 2003; Desmet & Declercq, 2006). We first present the results from two experiments run in the lab. We then present the results from a more rigorous third experiment, which was conducted partially in the lab and partially by using a web-based crowd-sourcing method. In both the lab and in web-based experiments, we find evidence for joint storage of syntactic and lexical information for frequent types of ORCs (i.e., pronominal ORCs and ORCs with *that* as a relative pronoun).

## Experiment 1

Experiment 1 was conducted to establish that ORCs elicit a priming effect (i.e., that they are produced with greater frequency after ORC primes than after other types of primes) and to test whether this effect might be greater for pronominal ORCs compared to definite-NP ORCs.

### Methods

**Participants** Thirty-two native English speakers were recruited from the MIT community. They were paid at a rate of $10/hour.

**Design and materials** The experiment had four prime conditions: a baseline prime (a definite NP); an SRC prime (a definite NP + the relative pronoun *who* + a transitive verb); a definite-NP ORC prime (a definite NP + *who* + the

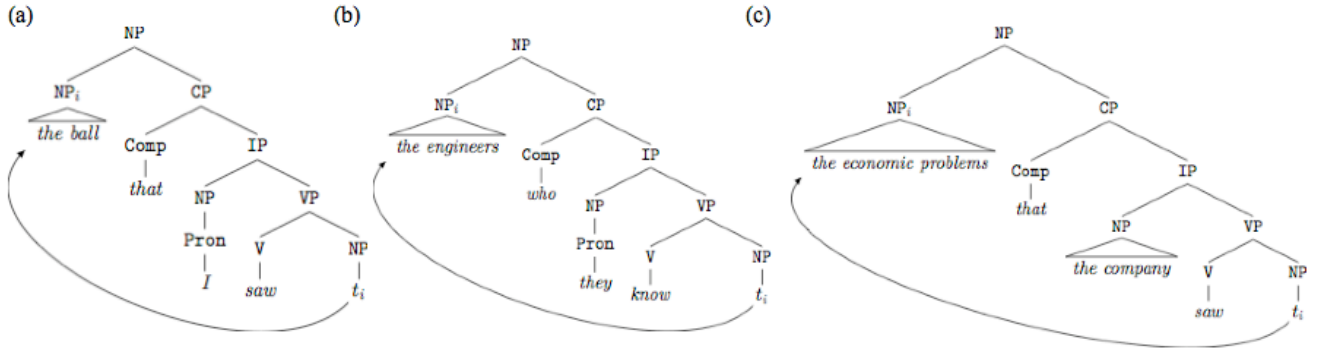


Figure 3: Schematic illustrations of the types of ORCs used in Experiments 1-3: pronominal ORCs (a-b) vs. definite-NP ORCs (c) and ORCs with "who" as the relative pronoun (b) vs. "that" as the relative pronoun (a and c)

determiner *the*); and a pronominal ORC prime (a definite NP + *who* + the personal pronoun *you* or *I*). Critically, in all but the baseline prime condition the participant was forced into an SRC or ORC completion. Each prime was followed by a target preamble consisting of a definite NP followed by *who*, in order to force a relative clause completion of the target preamble. A sample item is given in (1):

(1) a. *Baseline prime*
The screenwriter…
b. *SRC prime*
The screenwriter who noticed…
c. *Definite-NP ORC prime*
The screenwriter who the…
d. *Pronominal ORC prime*
The screenwriter who you / I…
e. *Target*
The marine who…

Twenty-four such items were created. Each participant only saw one version of each item (i.e., of each prime-target pair), according to a Latin Square design. Items were interleaved among 160 short filler preambles, in addition to 16 prime-target preamble pairs from an unrelated priming experiment, for a total of 240 sentences. Between three and five filler preambles appeared between each prime-target pair (following Desmet & Declercq, 2006). Subject to these constraints, four randomized lists of sentence preambles were created. Four additional lists were created by reversing the order of the trials in each list, for a total of eight lists distributed evenly among participants.

**Procedure** Participants were given a booklet containing the 240 sentence preambles and were instructed to form full sentences by completing each preamble with the first continuation that came to mind. They were instructed not to be too original or creative, but rather to be spontaneous and to write down the first thing they thought of. They were also prompted to move quickly through the experiment and to complete the sentence beginnings in the order that they were presented in. The experiment took between 45 minutes and an hour to complete.

### Results and Discussion

Prime-target pairs were excluded if either the prime preamble or the target preamble was left blank or not completed as a complete grammatical English sentence, affecting 15.36% of the data. We used the proportion of ORCs produced in the targets as the critical dependent measure. An overall SRC bias is to be expected in the RC completions given that SRCs are more frequent than ORCs (e.g., Roland et al., 2007). Indeed, participants overwhelmingly completed target preambles with SRCs; across conditions, only 2.77% of target preambles were completed with ORCs. The mean proportion and standard error of the mean for each condition, along with the relevant $\chi^2$ values and p-values for each proportion test, are displayed in Table 1. The proportion tests that reached significance ($p < .05$) are displayed in boldface, and those that approached significance ($p < .10$) are in italic boldface. Experiment 1 suggested priming of pronominal ORCs with the relative pronoun *who* (see Figure 4a). In Experiment 2, we attempted to replicate this finding with the relative pronoun *that*.

## Experiment 2

In Experiment 2, we had two goals. First, we wished to investigate the priming of ORCs containing *that* as relative pronoun. Second, we wanted to rule out a potential alternative explanation for the priming effect observed in Experiment 1. In particular, we wanted to test whether the presence of a pronoun in any RC, rather than in the subject position of an ORC specifically, might explain the priming effect. We therefore included a pronominal SRC condition, which contained a personal pronoun (*you* or *I*) in the embedded object position. Furthermore, we added another control (a complement clause prime condition) to test whether any embedded clause with a pronoun might be sufficient to elicit a greater proportion of ORC production in targets.

## Methods

**Participants** Twenty-nine native English speakers were recruited from the MIT community. They were paid at a rate of $10/hour.

**Design and materials** Given the extremely low base rate of ORC production across conditions in Experiment 1, no baseline condition was included in Experiment 2. We hypothesized that the complement clause prime condition would result in the lowest proportion of ORCs in the targets, and could thus serve as the new baseline condition. The experiment had five prime conditions: a definite-NP SRC prime (a definite NP + *that* + a transitive verb); a pronominal SRC prime (a definite NP + *that* + a transitive verb + the personal pronoun *you* or *me* + a preposition[1]); a complement clause prime (a definite NP + a sentential complement verb + *that* + the personal pronoun *you* or *I*); a definite-NP ORC prime (a definite NP + *that* + *the*); and a pronominal ORC prime (a definite NP + *that* + the personal pronoun *you* or *me*). As in experiment 1, each prime was followed by a target, as in (2f):

(2) a. *Complement clause prime (baseline)*
The screenwriter said that you / I…
b. *Definite-NP SRC prime*
The screenwriter that noticed…
c. *Pronominal SRC prime*
The screenwriter that noticed you / me by…
d. *Definite-NP ORC prime*
The screenwriter that the…
e. *Pronominal ORC prime*
The screenwriter that you / I…
f. *Target*
The marine that…

Thirty such items were created. Each participant saw one version of each item (i.e., of each prime-target pair), according to a Latin Square design. Items were interleaved among 120 short filler preambles, for a total of 180 sentences. Between three and five filler preambles intervened between each prime-target pair. Subject to these constraints, five randomized lists were created. Five additional lists were created by reversing the order of the trials in each list, for a total of ten lists distributed approximately evenly among participants.

**Procedure** The procedure was identical to that of Experiment 1.

## Results and Discussion

As in Experiment 1, only trials on which both the prime and target preambles were fully completed as intended were included in the analysis. This led to exclusion of 14.44% of the data. The base rate of ORC production (7.92%) was higher in Experiment 2 than in Experiment 1. The mean proportion and standard error of the mean for each condition, along with the relevant $\chi^2$ values and p-values for each proportion test, are displayed in Table 1.

In summary, we observed priming of ORCs in both pronominal and definite-NP ORC conditions in Experiment 2 (see Figure 4b). This is in contrast to Experiment 1, in which only pronominal ORCs elicited priming. Furthermore, rates of ORC production across conditions were higher in Experiment 2, plausibly due to the use of the

Table 1: Experiments 1-3 statistics

| Experiment Condition | mean | (SE) | $\chi^2$ | p-value |
|---|---|---|---|---|
| Exp. 1 (N=32) | .0277 | (.0064) | — | — |
| Baseline | .0183 | (.0105) | — | — |
| SRC | .0000 | (.0000) | 1.321 | .2504 |
| Def-NP ORC | .0248 | (.0123) | 0.001 | .9803 |
| ***Pronom. ORC*** | ***.0674*** | ***(.0197)*** | ***3.701*** | ***.0544*** |
| Exp. 2 (N=29) | | | — | — |
| Comp. clause | .0382 | (.0153) | — | — |
| Def-NP SRC | .0649 | (.0199) | 0.656 | .4181 |
| Pronom. SRC | .0411 | (.0165) | 0.000 | 1.000 |
| **Def-NP ORC** | **.1218** | **(.0263)** | **7.087** | **.0078** |
| **Pronom. ORC** | **.1274** | **(.0270)** | **6.344** | **.0118** |
| 3A-*who* (N=36) | .0286 | (.0052) | — | — |
| Def-NP baseline | .0181 | (.0104) | — | — |
| Pronom.baseline | .0179 | (.0102) | 0.000 | 1.000 |
| Def-NP comp | .0231 | (.0115) | 0.000 | 1.000 |
| Pronom. comp. | .0118 | (.0083) | 0.000 | .9840 |
| Def-NP ORC | .0585 | (.0180) | 2.699 | .1004 |
| Pronom. ORC | .0422 | (.0156) | 0.928 | .3354 |
| 3A-*that* (N=36) | .0964 | (.0093) | — | — |
| Def-NP baseline | .0769 | (.0206) | — | — |
| Pronom.baseline | .1176 | (.0248) | 1.170 | .2795 |
| Def-NP comp | .1124 | (.0243) | 0.863 | .3529 |
| Pronom. comp. | .0663 | (.0194) | 0.028 | .8679 |
| Def-NP ORC | .1176 | (.0248) | 1.170 | .2795 |
| Pronom. ORC | .0872 | (.0216) | 0.022 | .8818 |
| 3B-*who* (N=111) | .0411 | (.0037) | — | — |
| Def-NP baseline | .0308 | (.0078) | — | — |
| Pronom.baseline | .0287 | (.0076) | 0.000 | 1.000 |
| Def-NP comp | .0326 | (.0083) | 0.021 | .8849 |
| Pronom. comp. | .0288 | (.0076) | 0.000 | 1.000 |
| Def-NP ORC | .0459 | (.0098) | 1.456 | .2276 |
| **Pronom. ORC** | **.0809** | **(.0126)** | **11.54** | **.0007** |
| 3B-*that* (N=109) | .1457 | (.0064) | — | — |
| Def-NP baseline | .1275 | (.0144) | — | — |
| Pronom.baseline | .1292 | (.0144) | 0.000 | 1.000 |
| Def-NP comp | .1186 | (.0140) | 0.171 | .6790 |
| Pronom. comp. | .1621 | (.0158) | 2.004 | .1569 |
| Def-NP ORC | .1259 | (.0143) | 1.000 | .9368 |
| **Pronom. ORC** | **.1884** | **(.0169)** | **5.752** | **.0165** |

[1] The preposition was included so that participants would be required to produce some portion of the RC in all of the RC primes (a-d).

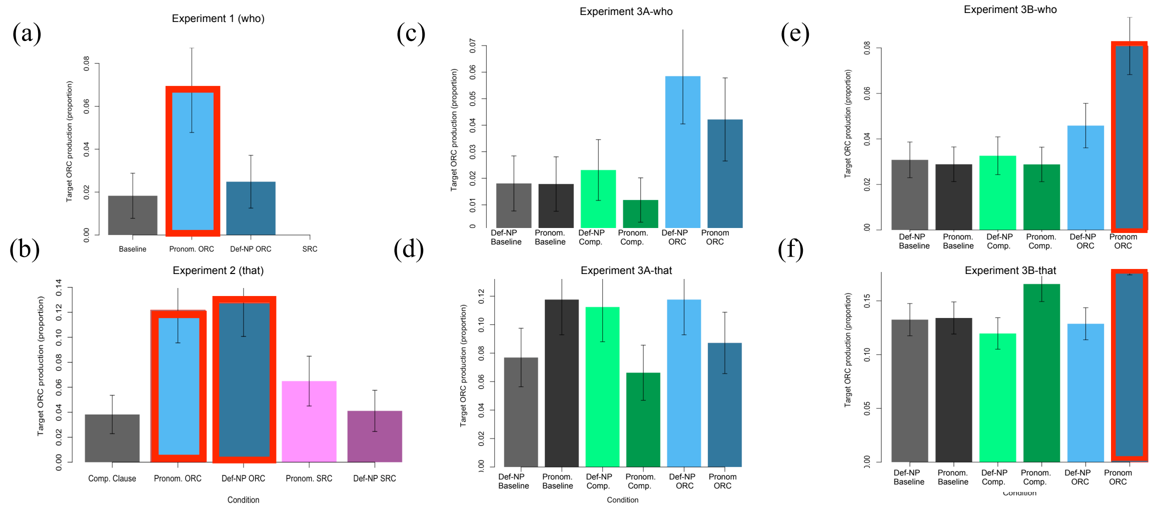


Figure 4: Results from Experiments 1 (a), 2 (b) and 3 (c-f). Mean proportion of ORC responses to targets per condition is plotted. In Experiment 1, there is a marginally significant effect of priming in pronominal ORCs. In Experiment 2, there is a significant priming effect in both definite-NP and pronominal ORCs. In Experiments 3B-*who* and 3B-*that*, there is a significant effect of priming in pronominal ORCs . Overall, the proportion of ORC responses in the *who* experiments is greater than the proportion of ORC responses in the *that* experiments.

relative pronoun *that* ($\chi^2$ = 16.846, p < .0001). In Experiment 3, we sought to replicate the effects observed in Experiments 1 and 2 including all the critical manipulations in the same design, increasing the number of participants by using a web-based crowd-sourcing service.

## Experiment 3

Experiment 3 used a fully crossed design, manipulating the NP type (full NP vs. pronoun) and structure of the prime (baseline vs. ORC vs. complement clause). In addition, relative pronoun (*who* vs. *that*) was manipulated between subjects. Given the small effect sizes in Experiments 1-2, Experiment 3 included a web-based component. In particular, we first conducted two experiments (with *who,* and with *that*) with similar numbers of subjects as in Experiments 1-2 in the lab (Experiment 3A-*who* and Experiment 3A-*that*). We then ran the same experiments using Amazon.com's Mechanical AMT service (henceforth referred to as "AMT") in order to a) gather data from a larger number of subjects, and b) validate the production priming method on AMT (see e.g., Munro, 2010, for validation of other experimental paradigms on AMT). These web-based experiments are referred to as Experiment 3B-*who* and Experiment 3B-*that*.

### Methods

**Participants** Thirty-six native English speakers recruited from the MIT community participated in Experiment 3A-*who*; 36 additional speakers from the same subject pool participated in Experiment 3A-*that*. These participants were paid at the rate of $10/hour for their time.

For the AMT studies, participants were excluded from analyses if they did not complete the entire experiment or if they were not native English speakers. This left 111 (out of 120) participants for Experiment 3B-*who* and 109 (out of 120) participants for Experiment 3B*that*. AMT participants were paid $1.25 for completing the task.

**Design and Materials** The design of Experiment 3 was as described above. A sample item is given in (3):

(3) a. *Definite-NP baseline prime*
The…
b. *Pronominal baseline prime*
You / I…
c. *Definite-NP ORC prime*
The screenwriter (who / that) the…
d. *Pronominal ORC prime*
The screenwriter (who / that) you / I…
e. *Definite-NP complement clause prime*
The screenwriter said that the…
f. *Pronominal complement clause prime*
The screenwriter said that you…
g. *Target*
The marine (who / that)…

Thirty such items were created, and 120 short filler preambles were also used in this experiment. Preambles were randomized as in Experiments 1-2.

**Procedure** For the two lab-based experiments, the procedure was identical to that of Experiments 1 and 2. For the AMT-based experiments, the procedure was also identical to the previous experiments, except that the preambles were presented electronically over Amazon.com Mechanical AMT's user interface, and subjects typed their responses rather than writing them by hand.

### Results and Discussion

As in Experiments 1 and 2, trials were excluded if either the prime or target preamble was not fully completed as intended. This led to exclusion of 6.2% (Experiment 3A-*who*), 3.2% (Experiment 3A-*that*), 13.8% (Experiment 3B-*who*), and 5.1% (Experiment 3B-*that*) of the data, respectively.

For each of the experiments, two-sample proportion tests were conducted for each condition against the definite-NP baseline condition from the relevant experiment. The mean proportion and standard error of the mean for each condition, along with the relevant $\chi^2$ values and p-values for each proportion test, are displayed in Table 2. The proportion tests that reached significance ($p < .05$) are displayed in boldface. In addition, the means and SEs for the average of ORC production for each experiment are shown.

As in Experiment 1, the two studies run on AMT (Experiment 3B) showed a significant effect of ORC priming, but only for the pronominal ORC (and not the definite-NP ORC) conditions (see Figure 4e-f). We did not observe an effect of ORC priming in Experiment 3A (see Figure 4c-d).

An additional two-sample proportion test was conducted across experiments to investigate the total number of ORC target completions following definite NP primes and pronominal primes. The mean proportion of ORC targets following definite-NP ORC primes was .0890 (SE = .0079), and the mean proportion of ORC targets following pronominal ORC primes was .1200 (SE = .0089). This difference was significant ($\chi^2$ = 6.434, p = .0112).

In addition, a large effect of relative pronoun was observed. Collapsing across all conditions of both of the *that* experiments, the mean proportion of ORC targets produced was .1334 (SE = .0053) compared to a mean of .0378 (SE = .0031) ORC targets in the *who* experiments. This effect was highly significant by a paired-sample proportion test, $\chi^2$ = 227.072, p < .0001.

## General Discussion

In this set of studies, we have provided the first evidence for priming of object relative clauses (ORCs). We have shown that the priming of ORCs is sensitive to the type of the NP in the embedded subject position (definite NP vs. pronoun) and the relative pronoun (*who* vs. *that*), two factors that have been also shown to affect (i) production frequencies, and (ii) processing complexity. These results are of special interest because an ORC is a complex and abstract syntactic structure involving an embedded clause, a displaced constituent, and a number of other syntactic positions (a possible representation is shown in Figure 2a). Although there are many accounts of syntactic priming (including transient activation, implicit learning, and pragmatic alignment; see Ferreira & Bock, 2006, for a recent review), all of them rely on the assumption that the primed element (or elements) must be stored in order to be primed. Our results imply that (some parts of) this complex ORC structure must be stored together with specific lexical items such as *that* or *you*. We recognize that some syntactic theories hold that *who* and *that* occupy different syntactic positions within an RC. Thus, it is possible that the abstract syntactic structures primed in the *who* and *that* conditions were not identical. Although this may complicate comparison between these two conditions, our results still indicate that some abstract syntactic structure was primed in both cases. We leave it to future work to consider this as a plausible explanation for the production frequency and priming differences put forth in this paper.

Our results cannot be reduced to the priming of a lexical element (i.e., the embedded subject pronoun or the relative pronoun) alone, as we do not observe increased production of ORCs in conditions with only a personal pronoun (e.g., the pronominal SRC, pronominal complement clause, and pronominal baseline conditions of Experiments 2-3). Furthermore, our results cannot be reduced to priming of a lexical item + embedded clause (e.g., *that* + embedded complement clause of any sort), as we do not observe increased production of ORCs after a complementizer and its complement clause (e.g., the "complementizer clause" conditions of Experiments 2-3). It appears that both the lexical element in a particular syntactic position and the abstract syntactic structure(s) underlying object relative clauses must be present for the priming to take place.

These results provide strong support for theories of syntactic structure which allow for the storage of a wide variety of both abstract and lexically-specific structures at different levels of granularity, even when the structures are fully compositional (e.g., Jackendoff, 2002; Goldberg, 2005). Of course, once the possibility of such storage is admitted, two natural questions arise: which structures are stored and why? One promising possibility is that the storage versus computation decision is the result of optimization of a tradeoff. Some tradeoff-based theories view the tradeoff in terms of computational resources; time spent computing can be reduced at the cost of space in memory and vice versa (e.g., Baayen, et al. 2007a). Under such theories high-frequency combinations of items may be stored as chunks in order to facilitate fast processing. Another recent proposal, O'Donnell (2011), takes a different view on the nature of the tradeoff. Under this approach, the tradeoff is viewed in terms of optimal predictions about productivity and reuse. The system optimizes its ability to predict future reuse of combinations of structure, while determining which parts of the system will be able to productively generate novel structures. We leave it to future work to explore these different possibilities.

## References


Alegre, M. and Gordon, P. (1999). Frequency effects and the representational status of regular inflections. Journal of Memory and Language, 40:41–61.

Arnon, I. and Snider, N. (2010). More than words: speakers are sensitive to the frequency of multi-word sequences. Journal of Memory and Language, 62:67–82.

Baayen, R.H., Dijkstra, T., & Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language*, *37*, 94-117.

Baayen, R. H., Burani, C., and Schreuder, R. (1997). Effects of semantic markedness in the processing of regular nominal singulars and plurals in Italian. In Booij, G. and van Marle, J., editors, Yearbook of morphology 1996. Kluwer Academic Publishers.

Bannard, C. and Matthews, D. (2008). Stored word sequences in language learning: The effect of familiarity on children's repetition of four-word combinations. Psychological Science.

Bertram, R., Laine, M., and Karvinen, K. (1999). The interplay of word formation type, affixal homonymy, and productivity in lexical processing: Evidence from a morphologically rich language. Journal of Psycholinguistic Research, 28(3):213–225.

Bloomfield, L. (1933). *Language*. New York: Holt, Rinehart and Winston.

Bock, J.K. Syntactic persistence in language production. (1986). *Cognitive Psychology*, *18*, 355-387.

Bod, R. (2001). Sentence memory: Storage vs. computation of frequent sentences. Talk presented at CUNY.

Bod, R., Scha, R., & Sima'an, K. (eds.) (2003). Data-Oriented Parsing. CSLI Publications, University of Chicago Press.

Chomsky, N. (1995). The Minimalist Program. MIT Press.

Desmet, T. & Declercq, M. (2006). Cross-linguistic priming of syntactic hierarchical configuration information. *Journal of Memory and Language*, *54*, 610-632.

Goldberg, A. E. (2005). Constructions at Work. Oxford University Press, Oxford.

Jackendoff, R. (2002). Foundations of Language. Oxford University Press, New York.

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

Marcus, M.P., Santorini, B., & Marcinkiewicz, M.A. (1995). *Treebank 2*. Linguistic Data Consortium, Philadelphia.

Mel'čuk, I. (1988). Dependency Syntax : Theory and Practice,. The SUNY Press, Albany, N.Y.

Munro, Robert. 2010. Crowdsourced translation for emergency response and beyond. *NSF Workshop on crowdsourcing and translation*, University of Maryland.

O'Donnell, T. J. (2011). Productivity and Reuse in Language. PhD thesis, Harvard University.

Reali, F. & Christiansen, M.H. (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 grammatical structures: A corpus analysis. *Journal of Memory and Language*, *57*, 348-379.

Scheepers, C. (2003). Syntactic priming of relative clause attachments: Persistence of structural configuration in sentence production. *Cognition*, *89*(3), 179-205.

Sag, I. A., Wasow, T., and Bender, E. M. (2003). Syntactic Theory: A Formal Introduction. CSLI, Stanford, CA, 2 edition.

Sereno, J. A. and Jongman, A. (1997). Processing of English inflectional morphology. Memory and Cognition, 25(4):415–437.

Steedman, M. (2000). The Syntactic Process. MIT Press, Cambridge, Ma.

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, 53*, 204-224.

Tremblay, A. (2009). Processing Advantages of Lexical Bundles: Evidence from Self-paced Reading, Word and Sentence Recall, and Free Recall with Event-related Brain Potential Recordings. PhD thesis, University of Alberta, Department of Linguistics.

Wanner, E. & Maratsos, M. (1978). An ATN approach to comprehension. In M. Halle, J. Bresnan, & G.A. 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.