

# Speakers encode silent structures: evidence from complementizer priming in English

Shota Momma<sup>1</sup>, Norvin Richards<sup>2</sup>, and Victor S. Ferreira<sup>3</sup>

<sup>1</sup>University of Massachusetts, Amherst

<sup>2</sup>Massachusetts Institute of Technology

<sup>3</sup>University of California, San Diego

Do speakers encode abstract structural representations devoid of perceptual-motor content? In six recall-based production experiments, we examined whether English speakers encode the null complementizer in sentence production using *structural priming*, the tendency for speakers to reuse the structure they have recently encountered. The results show that the null complementizer can be primed across distinct construction types and that this priming effect cannot be explained as the priming of the absence of the overt complementizer. These results are difficult to capture in semantic, pragmatic, or phonological terms. Furthermore, we evaluated transformer-based *Large Language Models* for their capacity to reproduce human priming patterns. Although they could reproduce basic priming effects, LLMs showed more sensitivity to constructional differences than humans, and they showed priming effects that were directionally opposite to human priming effects in some cases. This suggests that distributional cues alone are likely not sufficient for learning the abstract generalization governing the distribution of English complementizers. Based on these results, we argue that the structural representations speakers construct during production go beyond what they hear and say.

## Introduction

How do speakers represent sentences they utter? One view is that grammatical representations described in linguistic theories *are* mental representations that speakers construct during sentence comprehension and production. This view, which assumes the transparent relationship between theories of grammatical knowledge and syntactic processes in comprehension and production, is appealing when one attempts to integrate analytical linguistics and psycholinguistics (Lewis & Phillips, 2015; Momma & Phillips, 2018). However, theories of grammatical knowledge, especially those in the generative grammar tradition, may appear difficult to use as the representational basis of perception and production theories. These theories routinely posit structures that do not have direct perceptual or motor content, which we collectively call *null elements*. If something has no perceptual-motor content, why should we assume its relevance in perception and production?

In this article, we investigate if speakers encode null ele-

ments, focusing on the *null complementizer*, a type of phonologically empty functional projection assumed to introduce a clause in sentences like *the girl believes 0 the boy has a cat* and *I love the cat 0 the boy has* (here, 0 represents the null complementizer). We evaluate whether speakers encode the null complementizer using *structural priming* (Bock 1986; Levelt and Kelter 1982; see Pickering and Ferreira 2008 for an overview and Mahowald, James, Futrell, and Gibson 2016 for a meta-analysis), the tendency for speakers to repeat the structures they have recently encountered.

We report six recall-based structural priming experiments in English testing whether speakers encode the null complementizer in various types of embedded clauses. Based on the results of these experiments, we argue that speakers indeed encode the abstract null complementizer shared across different types of embedded clauses as a cognitively active but phonologically empty representation. This conclusion has broad implications for the nature of grammatical knowledge and how grammatical knowledge is put to use as we discuss below.

## Null elements in linguistics and psycholinguistics

Linguistic theories, especially those in the mainstream generative tradition, posit various types of null elements in syntax. Examples include *traces* or *copies* that record information about the underlying position of a displaced constituent (e.g., the null element, represented by *e* in sentences like *Who does the girl believe the boy met e* ?), null op-

---

The authors acknowledge the support from an NSF grant BCS-2234229 to SM. Correspondence concerning this article should be addressed to Shota Momma, Department of Linguistics, University of Massachusetts, Amherst, N408 Integrative Learning Center 650 North Pleasant Street Amherst, MA 01003. E-mail: snegishi@umass.edu

erators (Browning, 1987; Chomsky, 1981; Contreras, 1984; Haik, 1985), implicit arguments (Bhatt & Pancheva, 2006; Roeper, 1987), complex syntactic structures inside ellipsis sites (Merchant, 2001; Sag, 1976; Williams, 1977), and various null functional categories including the null complementizer (Bošković & Lasnik, 2003; Kishimoto, 2006; Pesetsky, 1995; Stowell, 1981), voice heads (Kratzer, 1996), determiner heads (Abney, 1987) among others (see Rizzi and Cinque 2016). Null elements are a standard theoretical device for grammatical analyses in the mainstream generative tradition.

Why null elements? Null elements play critical roles in formulating abstract generalizations as well as making abstract distinctions between superficially similar constructions. For example, in the Government and Binding theory (Chomsky, 1981), traces allow constraints on phrasal movement (the dislocation of phrases from their canonical position, as in *wh*-sentences) to be subsumed under the binding theory, which also explains the distribution of various types of noun phrase. Null elements also allow structural distinction between seemingly similar constructions like *the cat seems to be loved* (a so-called *raising* structure) and *the cat wants to be loved* (a so-called *control* structure) (Polinsky 2013; Rosenbaum 1965, cf. Hornstein 1999) and critical in explaining various differing behaviors of those constructions. In addition, null elements allow the syntax-semantics interface to be simpler. For instance, if one assumes that no structure exists in ellipsis sites (e.g., *e* in *Mary fed the dog, and John did e, too*), a rule for mapping elliptical expressions to interpretation must become more complicated (see, e.g., Aelbrecht 2010).

Despite their theoretical utility, various (psycho-)linguistic theories eschew null elements. In analytical linguistics, some theories are sometimes dubbed as *WYSIWYG* (*What You See Is What You Get*) theories of grammar. These theories include a certain subclass of *Combinatory Categorical Grammar* (CCG, Steedman 2001 cf. the *Principle of Adjacency*), *Simpler Syntax* (Culicover and Jackendoff 2006, cf. the *simpler syntax hypothesis*), and constructionist approaches (see, e.g., Goldberg and Suttle 2010). These theories adopt an explicit statement against positing null elements as a guiding principle in theory development. Other theories, including *Generalized Phrase Structure Grammar* (GPSG, Gazdar 1981), *Head-driven Phrase Structure Grammar* (HPSG, Pollard and Sag 1994, Ginzburg and Sag 2000), and *Lexical Functional Grammar* (LFG, Bresnan, Asudeh, Toivonen, and Wechsler 2015; Kaplan and Bresnan 1981) generally avoid introducing null elements with limited exceptions. A prominent version of Dependency Grammar (De Marneffe & Nivre, 2019) uses null elements sparingly, in fact, in only one construction (the gapping construction where a predicate is missing in a clause). Overall, many grammatical theories avoid positing

null elements as theoretical constructs.

A parallel controversy exists in psycholinguistics. Some have argued against the existence of null elements, in particular, the existence of traces or invisible copies (see, e.g., Culicover and Jackendoff 2006; Pickering 1993; Sag and Fodor 1994) primarily based on simplicity grounds (though see Pickering and Barry 1991 for providing non-simplicity based argument, but see also Gibson and Hickok 1993 and Gorrell 1993 for criticisms). Others have pointed out that perception routinely goes beyond distal stimuli in cognition generally (e.g., the Kanizsa triangle: Kanizsa et al. 1955) and argue that there is nothing unnatural about positing null elements in theories of linguistic perception (Phillips & Wagers, 2007).

In this article, we focus on whether speakers encode the *null complementizer*, like in *The girl believes 0 the boy has a cat* in sentence production (0 represents the null complementizer that can occur in place of the overt complementizer *that*). The null complementizer is a subtype of phonologically null functional category posited in some (e.g., Bošković and Lasnik 2003; Chomsky 1981, 1995; Kishimoto 2006; Stowell 1981) but not all (e.g., Müller, Abeillé, Borsley, and Koenig 2021, Steedman 2001) syntactic theories. Not surprisingly, those theories that assume no null complementizers are theories that generally avoids null elements. The question of the operation of the null complementizer relates to broader questions such as whether our grammatical knowledge involves abstraction that goes beyond sensory-motor experiences. Hence, studying the null complementizer has general implications for theories of grammar and cognition.

### Structural priming and complementizer omission

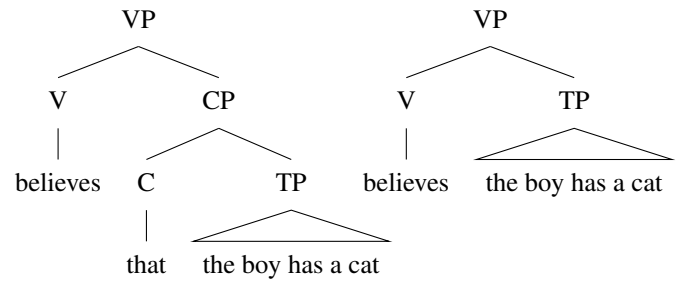
One of the most well-studied phenomena in sentence production research is *syntactic* or *structural priming* (Bock, 1986; Levelt & Kelter, 1982), the tendency for speakers to reuse the structure of sentences they have recently encountered. A classic example involves the dative alternation in English. Bock (1986) showed that after repeating a *prepositional dative* (PD) prime like *The rockstar sold some cocaine to an undercover agent*, speakers are more likely to reuse the PD structure again, for example, *the man is reading a story to the boy*. No word overlap between prime and target sentences is necessary for structural priming (Bock, 1986, 1989), though repeating a word, especially the head of the primed structure, enhances the priming effect (Pickering & Branigan, 1998). The priming effect has been shown to be *bidirectional*, meaning that PD and DO can each be primed compared to a neutral baseline (Pickering & Branigan, 1998). Structural priming has proven to be a useful tool for investigating the structural representations speakers construct during production (Branigan & Pickering, 2017; Pickering & Ferreira, 2008) and comprehension, to the extent the prim-

ing occurs from comprehension to production (Bock, Dell, Chang, & Onishi, 2007).

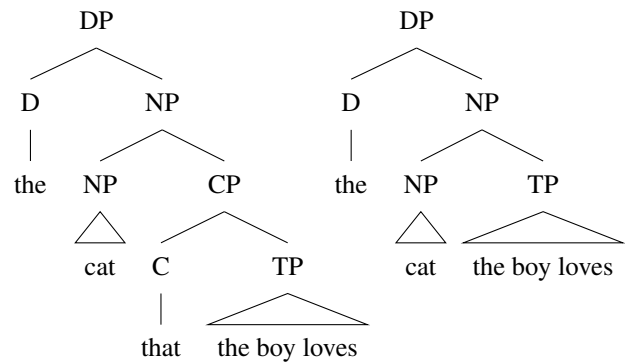
The most relevant type of structural priming for current purposes is the priming of English complementizers. Complementizers are, roughly, lexical items that introduce a clause, like *that* in *the girl believes that the boy has a cat*. In some, but not all, syntactic theories, the complementizer is hypothesized to exist even when there is no phonology associated with it.

In a series of recall-based production studies, V. S. Ferreira (2003) showed that the *lack* of the overt complementizer *that* can be primed. V. S. Ferreira (2003) showed that, the probability of *that* production in target sentences like *the teacher noticed (that) the failing student skipped the class* was lower given prime sentences without *that* like *the director announced Hollywood's hottest actor would be playing the part* than given neutral prime sentences, like *The director enticed Hollywood's hottest actor to play the part*. Building on Ferreira (2003), Jaeger and Snider (2008) found the complementizer priming effect using naturalistic corpora. They found that the overt complementizer *that* as it occurs in a VP-complement clause can be primed in a dialogue. They also extended the prior studies by showing that *that* as it occurs in relative clauses (e.g., *the boy (that) the girl met*) can also be primed. Interestingly, unlike V. S. Ferreira (2003), Jaeger and Snider (2008) showed the priming of the overt complementizer as opposed to the priming of the null complementizer. They suggested that this difference was due to the difference in which structure is more surprising given a specific circumstance and that more surprising structures show bigger priming effects.

Critically, the results from neither V. S. Ferreira (2003) nor Jaeger and Snider (2008) necessarily suggest that speakers encode the null complementizer beyond the mere absence of the overt complementizer *that*. For instance, the complementizer omission effect can be explained as resulting from the priming of the phrase structure rule:  $VP \rightarrow V TP$  (for complement clauses) or  $NP \rightarrow NP TP$  (for relative clauses). TP is a shorthand for Tense Phrase, a type of constituent that includes all but the complementizer of a clause in a standard analysis. This account falls under the standard *configurational* account of structural priming (e.g., Bock and Loebell 1990; Pickering and Branigan 1998; Pickering and Ferreira 2008), which suggests that structural priming results from the priming of a phrase structure configuration. Likewise, the priming of the overt complementizer *that* found in Jaeger and Snider (2008) can be captured as the priming of the phrase structure rule:  $VP \rightarrow V CP$ , assuming the structure like those in below.



The same holds for relative clause complementizer priming found in Jaeger and Snider (2008). Assuming the structures like below, the complementizer omission priming (though it was not observed in Jaeger and Snider 2008) can be construed as the priming of the phrase structure rule:  $NP \rightarrow NP TP$ , and the overt complementizer priming can be construed as the priming of the phrase structure rule:  $NP \rightarrow NP CP$ .



Thus, the extant evidence can be captured by the standard configurational account of syntactic priming assuming no null complementizer.

### Current study

The current study aims to assess whether complementizer priming can be explained without assuming the null complementizer. The configurational account assuming no null complementizer has difficulty predicting the null complementizer priming from relative clauses to VP-complement clauses. This is because VP complement clauses and relative clauses do not share the relevant phrase structure configuration if no null elements are assumed to exist. The configurational account assuming no null complementizer must assume that the rule like  $VP \rightarrow V TP$  is what primes the complementizer omission observed by V. S. Ferreira (2003). However, relative clauses are not an argument of a verb and thus cannot involve a phrase structure rule like  $VP \rightarrow V TP$ , as can be seen in the tree for the relative clause containing the noun phrase above. Thus, the configurational account assuming no null complementizer predicts that relative clause complementizers and VP-complement clause complementizers should not prime each other. Experiments 1 and 2 test this prediction.

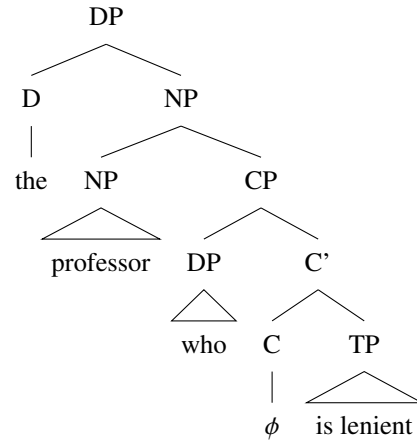
If the null complementizer is assumed to exist, cross-constructional null complementizer priming is predicted to be observed. There are at least two accounts that would predict null complementizer priming across constructions. The first account is a version of the configurational account, but it assumes the null complementizer. Under this account, the priming of complementizer omission between VP-complement and relative clauses can be captured as the priming of the phrase structure rule like  $CP \rightarrow C_{null} TP$ , which is involved in both VP-complement clauses and relative clauses with no overt complementizer. Critically, this account is possible only if the null complementizer is assumed to exist.

The second account is the *functional category-based account* of structural priming proposed by Momma (2022). Momma (2022) suggested that the majority of structural priming effects can be captured by assuming that structural priming results from the strengthening of the mapping between a message component and a (potentially phonologically null) syntactic head. Critically, this account is only possible if various null functional categories, such as the *applicative* head in double object dative structures, the *voice* head encoding passive voice, and the null complementizer head, are assumed. Under this account, dative priming can be thought of as the strengthening of the mapping between the relevant message feature and either the applicative head or the abstract preposition (encompassing both *for* and *to*; Bock 1989) associated with the *goal* or *recipient* role, presumably represented at the message-level. Passive priming can be thought of as the strengthening of the mapping between the relevant message feature (like agentivity) and either active or passive voice head (or possibly, the preposition *by*, Bock and Loebell 1990; Ziegler, Bencini, Goldberg, and Snedeker 2019). Complementizer priming can be thought of as the strengthening of the mapping between the relevant message feature (like embedded proposition) and either the overt complementizer or the null complementizer.

The two accounts assuming the null complementizer both make three untested predictions. First, the VP-complement clauses and relative clauses share the relevant phrase structure rule ( $CP \rightarrow C_{null} TP$ ) or functional category (the null complementizer). Therefore, unlike the configurational account assuming no null complementizer, the accounts assuming the null complementizer predict that the priming of complementizer omission should occur even when the null complementizer occurs in two distinct constructions, such as between VP-complement clauses and relative clauses. Note that the current study does not aim to distinguish between these two accounts assuming the null complementizer. The primary goal instead is to assess if speakers encode the null complementizer.

Second, when combined with contemporary syntactic theories of clause structures, the accounts assuming the null complementizer predict that the relative clauses with the rel-

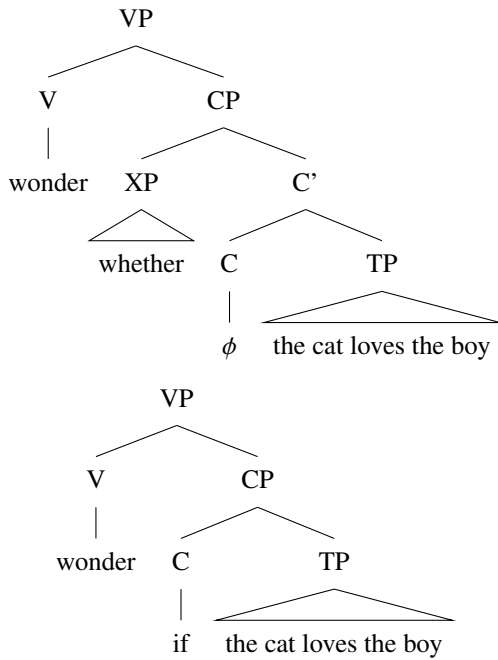
ative pronoun *who* should prime the null complementizer compared to the neutral baseline. The reason is that relative clauses with *who*, like *the professor who is lenient* have the following syntactic structure, according to a common assumption (see Bhatt 2002 for an overview of relative clause analyses).



Under this analysis, the relative pronoun *who* (as well as *which* and *what*) does not occupy the complementizer position. Instead, it occupies the specifier position of the embedded CP. The complementizer position is instead occupied by the null complementizer, just like in relative clauses with no relative pronoun or overt complementizer, like *the cat the boy loves*. Note that an expression like *the professor who that is lenient*, where the complementizer *that* and the relative pronoun *who* occur, is blocked by an independent constraint dubbed as the *doubly filled comp filter* (Chomsky & Lasnik, 1977). If this analysis is correct, the accounts assuming the null complementizer predict that relative clauses with a relative pronoun should prime the null complementizer in VP-complement clauses. Experiments 3, 4, and 5 tested this prediction.

Finally, some analyses of syntax suggest that interrogative VP-complement clauses containing *whether* (e.g., *I wonder whether the cat loves the boy*), but not minimally different clauses containing *if* (e.g., *I wonder if the cat loves the boy*), may contain the null complementizer (Kayne, 1991; Larson, 1985; Wu, 2022). The reason is that *whether* can be analyzed as a wh-phrase that resides in the specifier of the embedded CP, and *if* can be analyzed as occupying the complementizer position. The empirical reason for analyzing *whether* as forming a wh-phrase as opposed to being a complementizer is indirect. However, one piece of evidence is the cross-linguistic correlation between the presence or absence of *either* and *whether*. Languages that have (the translation equivalent of) *either* in their vocabulary (English and Polish, Wu 2022) also have *whether*, and languages that do not have *either* in their vocabulary (e.g., French and Bengali, Kayne 1991; Wu 2022) also do not have *whether*. Another

piece of evidence is that, just like *either*, *whether* but not *if* can optionally co-occur with *or not* (in English). These observations, combined with the fact that both *either* and *whether* express disjunction, suggest that *whether* may be the wh version of *either*. Under this analysis, *whether* resides in the same position as other wh-phrases as a result of wh-movement, that is, the specifier position of the embedded CP. If this analysis is correct, verb phrases like *wonder iff/whether the cat loves the boy* have the following syntactic structures (given the uncertainty about the lexical category of *whether*, I use XP for the phrase headed by it).



Under this analysis, the accounts assuming the null complementizer predict that *whether*-clauses prime the null complementizer (to the extent complementizers that occur in interrogative clauses prime those in non-interrogative clauses). Experiment 6 tested this prediction.

### Experiment 1

Experiment 1 aims to test if complementizer priming occurs between VP-complement clauses and relative clauses. Speakers memorized the target and prime sentences (in this order) like the following and recalled target sentences:

- (1) The director announced (that) the actor would be in the new movie. [Target]
- (2) a. The professor thought (that) the students appreciated the idea during the class. [VP-complement prime]  
 b. The professor appreciated the thoughts (that) the students expressed during the class. [Relative clause prime]

In relative clause primes, the noun a relative clause modifies is always a nominal counter-part of clause-introducing verbs (e.g., *thought*, *claim*, *guess*), etc. This was to make the sentence as similar as possible across conditions.

### Method

#### Participants

48 self-reported monolingual speakers of English were recruited via Prolific Academic (the sample size was the same as V. S. Ferreira 2003). Each participant was paid \$5 for an approximately 25-30 minute session. Two participants were replaced because they did not follow instructions or their recordings were unintelligible.

#### Materials

24 pairs of prime and target sentences like (1), (2a), and (2b) were created. See the appendix for the full list of stimuli. All sentences were bi-clausal sentences with clause-embedding verbs. There were 12 distinct clause-embedding verbs, each repeated two times in target sentences. The embedding verbs were never repeated within each pair.

In experimental trials, recall cues consisted of a matrix subject noun (e.g., *student*), an embedding verb (e.g., *said*), an embedded subject noun (e.g., *teacher*), and an embedded predicate (e.g., *nice*). The order of these cues was randomized across trials.

There were a total of 32 filler sentence pairs (a randomly chosen example pair is *The man wanted to understand the theory, but he couldn't.* and *The hobbyist maintained that the antique cars were one of a kind.*), 8 of which were designed to look similar to experimental sentence pairs. In 14 of these filler sentence pairs, recall cues corresponded to the second sentence. Thus, participants were asked to recall the first sentence 25% of the time. This ensured that speakers paid attention to the second sentence, which served as the prime sentence in critical trials. The full list of experimental and filler sentences is available in the OSF repository (MXXX).

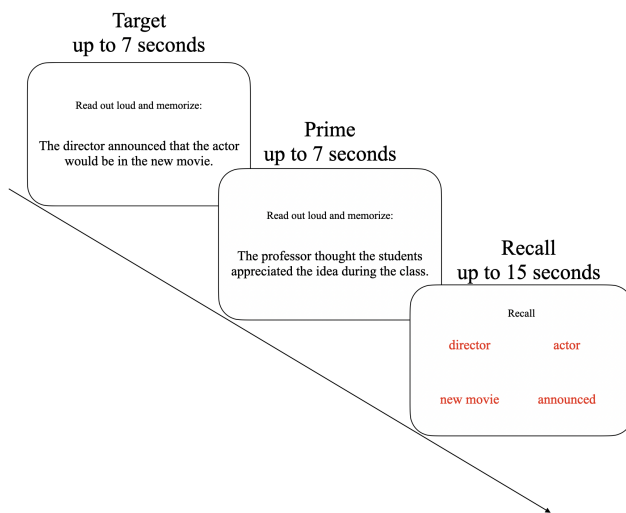
#### Procedure

The experiment was conducted online using Penn Controller Ibex (Zehr & Schwarz, 2018). Participants were first given the following instructions.

*In each trial in this experiment, you will see two sentences, one at a time. Your task is to read out loud and memorize the two sentences for later recall. You are given 7 seconds to read out loud and memorize each sentence. After you read and memorize two sentences, you will be prompted to recall one of them. The prompt will be words and phrases that appeared in one of the two sentences. For example, if you memorize 'The girl said that the boy went to the store' and 'The breeder believes that the dog chased the cat', you may see 'the girl said that, 'store', 'boy', 'went', (the order is*

random), as the prompt to recall the sentence 'The girl said that the boy went to the store.' Please recite the sentence that includes all the prompt words or phrases aloud.

Following this instruction, participants went through two practice trials, which had the same task structure as the experimental trials. Following the practice trials, the main experimental trials began, which had the following sequence of events. First, participants were asked to press the space key to start each trial. Immediately after the key press, a sentence (a target sentence in critical trials) was presented on the screen. After 7 seconds or when the space key was pressed, whichever is earlier, the first sentence was removed from the screen and another sentence (a prime sentence in critical trials) was presented. After 7 seconds or when the space key was pressed, whichever is earlier, the second sentence was removed and recall cues were presented, each word occupying one of the four positions on the screen (see Figure 1). The spatial positions of four cue words/phrases were random across trials. Participants were given up to 15 seconds to recall the sentence. Recall cues were kept on the screen until the time limit or until the space key was pressed. Participants' voice responses were recorded using participants' microphones. The schematic illustration of the task structure is shown in Figure 1.



**Figure 1**

*A schematic illustration of the task.*

Experimental stimuli were distributed according to a Latin Square design, and trials were fully randomized for each participant. In all subsequent experiments (Experiments 2-5), the experimental procedure was identical.

### Scoring and analysis

Audio recordings of participants' responses were coded by the first author without the knowledge of experimen-

tal condition. Target productions were coded for whether they contained sentential complements that in principle allow *that*; any trials in which the target did not contain sentential complement or any trials in which the target contained lexical items that prevent the use of *that* (e.g., *whether*, *if*, a *wh*-phrase at the beginning of complement clauses) were excluded from the main analysis. Responses that didn't include the target verb (always provided as a memory cue) were also excluded from the analysis.

Among all analyzable responses, those that contained *that* were coded as 1, and those that contained no overt complementizer were coded as 0.

The scored responses were then analyzed using Bayesian hierarchical mixed-effects regression assuming a Bernoulli distribution with the logit link function, with subject and item as random effects and with maximal random effects structure (Barr, Levy, Scheepers, & Tily, 2013). The fixed effects were *Prime Type* (whether the prime contained the null or the *that* complementizer; null vs. *that*, coded as -0.5 and 0.5), *Clause Type* (whether the prime contained a relative clause or a VP-complement clause, coded as -0.5 and 0.5), *Target Type* (whether the prime contained the null or the *that* complementizer, coded as -0.5 and 0.5 for crossed models), and interactions between them. All analyses were conducted using R (R Core Team, 2021) and the *brms* package (Bürkner, 2021). For each model, we ran four Monte Carlo Markov Chains with 5000 samples. The first 2000 samples were always discarded. All R-hat values were 1.

For all Bayesian mixed-effects analyses, priors were defined as Gaussian distribution with a mean of 0 and a SD of 1.5 (for the intercept) and 1 (for all fixed effects). We also used LKJ regularizing prior (Lewandowski, Kurowicka, & Joe, 2009) with  $\eta = 2$ . For all analyses, 95% credible intervals (based on percentiles) were used to make inferences. That is, we consider an effect to be present when the 95% credible intervals do not include 0.

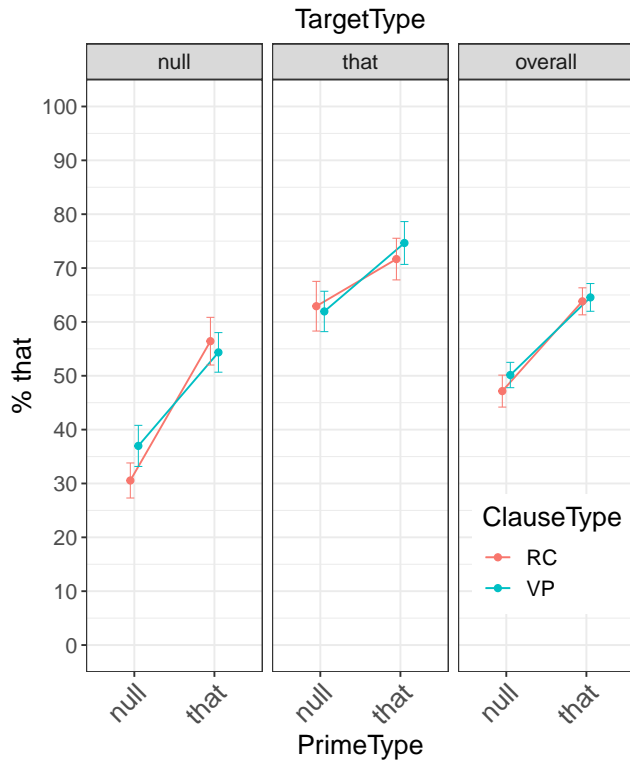
### Results

22 out of 1152 trials (1.9 %) did not meet the scoring criteria discussed above, so they were removed from further analysis. Figure 2 shows the average rate of *that* production in each condition, and Figure 3 shows the posterior estimates of fixed effect parameters.

When prime sentences contained VP-complement clauses, speakers on average produced the *that* complementizer 64.5% of the time given *that* complementizer primes. In contrast, they on average produced the *that* complementizer 50.0% of the time given null complementizer primes. That is, there was a 14.5% complementizer priming effect when the prime sentences contained VP-complement clauses. When prime sentences contained relative clauses, speakers on average produced the *that* complementizer 64.1% of the time given *that* complementizer primes. In contrast, they

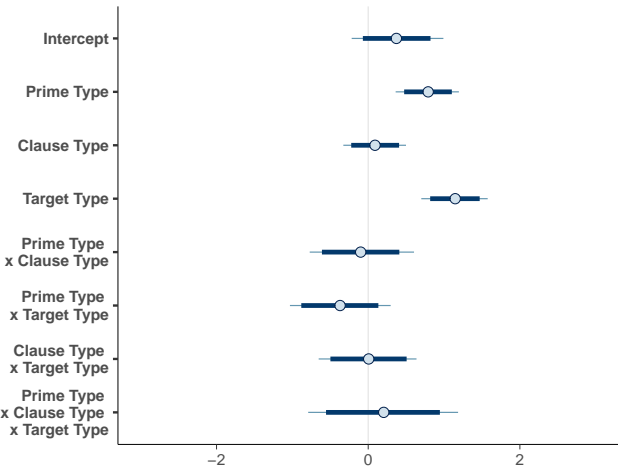
on average produced the *that* complementizer 48.2% of the time given null complementizer primes. That is, there was a 15.9% complementizer priming effect when the prime sentences contained relative clauses.

As can be seen in Figure 3, there was strong evidence for the main effect of Prime Type ( $\hat{\beta} = 0.79$ ,  $SE = 0.16$ , 95% CrI = [0.47, 1.11]), with no interaction between Prime Type and Clause Type ( $\hat{\beta} = -0.10$ ,  $SE = 0.26$ , 95% CrI = [-0.61, 0.41]). This suggests that the complementizers in VP-complement clauses and relative clauses primed complementizers in VP-complement clauses to a similar degree. If anything, the priming effect was numerically (1.4%) stronger from relative clauses to VP-complement clauses than from VP-complement clauses to VP-complement clauses. There was also evidence for the main effect of Target Type ( $\hat{\beta} = 1.15$ ,  $SE = 0.17$ , 95% CrI = [0.82, 1.47]), simply suggesting that speakers were more likely to use *that* when the original target sentences contained it. There was no reliable evidence for any other effects.



**Figure 2**

*By-condition proportions of that use in Experiment 1. Error bars represent within-subject standard errors of by-subject means.*



**Figure 3**

*Posterior estimates of the fixed effects with 95% (thick lines) and 99% (thin lines) credible intervals in Experiment 1.*

## Discussion

In Experiment 1, we observed the complementizer priming from relative clauses to VP-complement clauses. To the extent this effect reflects the priming of complementizer omission (as opposed to the priming of the overt complementizer use), it cannot be explained by the configurational account assuming no null complementizer. However, the implications of the current results depend on the directionality of the observed effect. Because there was no neutral condition, there is no way to tell if the effect was simply the priming of *that*, the priming of the null complementizer, or both. If the current results are simply the priming of *that* like in Jaeger and Snider (2008) as opposed to the priming of the null complementizer, they are still compatible with the configurational account without assuming the null complementizer. Experiment 2 aims to assess the directionality of the complementizer effect in the current experimental context.

## Experiment 2

In Experiment 1, there was complementizer priming across constructions, but the results did not strictly establish that the null complementizer, as opposed to the overt complementizer, can be primed across constructions. This was because the observed priming effect can be due entirely to the priming of the *that* complementizer (cf. Jaeger and Snider 2008). To address this issue, Experiment 2 examined the directionality of the cross-constructional complementizer priming by comparing prime sentences like the following:

- (3) a. The professor appreciated the thought (that / 0) the student expressed. [Complementizer prime]

- b. The professor appreciated the thought of that / the student. [Demonstrative prime]

Sentences like (3a) contained a relative clause with overt or null complementizers. In contrast, sentences like (3b) contained no embedded clauses but contained the demonstrative *that* or not. By comparing null complementizer prime sentences and prime sentences that contained neither embedded clauses nor the phonological form *that*, we can assess whether the null complementizer can be primed above and beyond the mere absence of the phonological form *that*.

If the cross-constructional null complementizer priming is observed, the result would be difficult to capture under the configuration account assuming no null complementizer because the shared phrase structure rule between VP-complement clauses and relative clauses cannot be stated without assuming the null complementizer. In addition, such a result would suggest that the null complementizer priming cannot be reduced to the priming of the mere absence of *that* and instead suggest that the null complementizer is represented as a primable entity.

## Methods

### Participants

48 self-reported monolingual speakers of English were recruited via Prolific Academic. Each participant was paid \$5 for an approximately 25-30 minute session. One participant was replaced because they did not follow instructions or their recordings were unintelligible.

### Materials

The target sentences were identical to those used in Experiment 1. The prime sentences were similar to those used in the relative clause conditions in Experiment 1, the syntactic category of *that* was manipulated. It functioned either as the complementizer *that* (as in 3a) or the demonstrative *that* (as in 3b). The demonstrative *that* was contrasted with the determiner *the*. Within each item set, the clause-taking verbs and the nouns that relative clauses modify in the prime sentences shared the same root (e.g., *claim* as a noun and a verb).

### Scoring and analysis

The same scoring criterion as in Experiment 1 was used. The analysis procedure was also the same, except that the new factor *Category* (demonstrative vs. complementizer, coded as -0.5 and 0.5, respectively) replaced *Clause Type*.

## Results

50 out of 1152 trials (4.3%) did not meet the scoring criteria described above, so they were removed from further analyses.

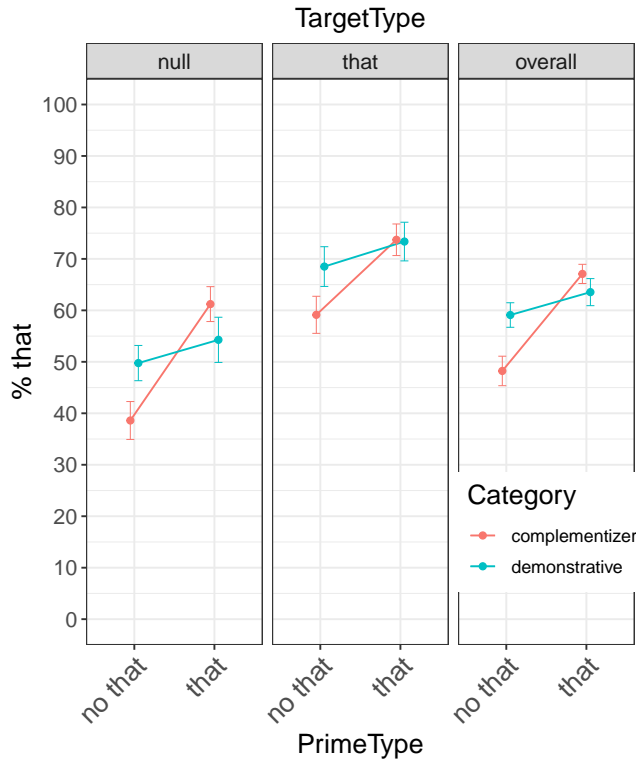
Figure 4 shows the average rate of *that* production in each condition. When prime sentences contained relative clauses (i.e., when *Category* was *complementizer*), speakers on average produced the *that* complementizer 67.1% of the time given *that* complementizer primes. In contrast, they produced the *that* complementizer 48.2% of the time given null complementizer primes. That is, there was an 18.9% complementizer priming effect when the prime sentences contained relative clauses. When prime sentences contained no relative clauses (i.e., when the *Category* was *demonstrative*), speakers on average produced the *that* complementizer 63.5% of the time given *that* complementizer primes. In contrast, they on average produced the *that* complementizer 59.1% of the time given null complementizer primes. That is, there was a 4.2% complementizer priming effect when the prime sentences contained no relative clauses.

As can be seen in Figure 5, there was strong evidence for the main effect of Prime Type ( $\beta = 0.65$ , SE = 0.17, 95% CrI = [0.31, 0.99]), but this was qualified by the interaction between Prime Type and Category ( $\beta = -0.66$ , SE = 0.28, 95% CrI = [-1.19, -0.12]). Further analyses using nested models suggest that there was evidence for the complementizer priming only in the complementizer conditions ( $\beta = 0.89$ , SE = 0.21, 95% CrI = [0.47, 1.30]) but not in the demonstrative conditions ( $\beta = 0.33$ , SE = 0.20, 95% CrI = [-0.07, 0.73]). There was also evidence for the main effect of Target Type ( $\beta = 0.98$ , SE = 0.17, 95% CrI = [0.65, 1.32]). But this simply suggests that speakers were more like to use *that* when the original target sentences contained it. There was no reliable evidence for any other effects.

In addition, we assessed the simple effect of Category within each level of Prime Type, to assess whether the null complementizer primes the null complementizer above and beyond the absence of *that* and whether the *that* complementizer primes the *that* complementizer above and beyond the mere presence of homophonous (demonstrative) *that*. That is, we compared the conditions where prime sentences were like *The professor appreciated the thoughts that the student expressed* and *The professor appreciated the thoughts of that student*, and the conditions where prime sentences were like *The professor appreciated the thoughts the student expressed* and *The professor appreciated the thoughts of the student*.

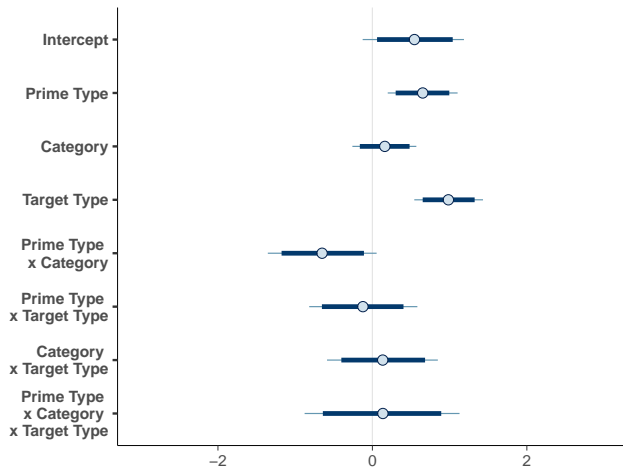
These analyses revealed evidence that the null complementizer primed the null complementizer above and beyond the mere absence of *that* ( $\beta = 0.43$ , SE = 0.20, 95% CrI = [0.05, 0.82]). There was little to no evidence for *that* complementizer priming the *that* complementizer above and beyond the demonstrative *that* ( $\beta = -0.13$ , SE = 0.20, 95% CrI = [-0.51, 0.25]), like in V. S. Ferreira (2003) but unlike in Jaeger and Snider (2008). These results suggest that the cross-constructional null complementizer priming effect observed here cannot be explained by the mere absence of the phonological form *that*.





**Figure 4**

By-condition proportions of *that* use in Experiment 2. Error bars represent within-subject standard errors of by-subject means.



**Figure 5**

Posterior estimates of the fixed effects with 95% (thick lines) and 99% (thin lines) credible intervals in Experiment 2.

## Discussion

Experiment 2 showed that the null complementizer can indeed be primed across constructions above and beyond the

mere absence of the phonological form *that*. Because the null complementizer was primed across constructions (from relative clauses to VP-complement clauses), the configurational account assuming no null complementizer would struggle to capture the current results.

## Experiment 3

Experiment 3 examined the further prediction of the accounts positing the null complementizer. Specifically, modern syntactic theories in the mainstream generative tradition suggest that relative clauses contain the null complementizer even when *that* is not optional, like when a relative pronoun like *who* or *which* occupies the left edge of a relative clause (e.g., *the professor who is lenient*). Under this analysis, a relative clause with *who* can prime the null complementizer in a VP-complement clause. Experiment 3 tested this prediction using prime sentences like the following.

- (4) a. They appreciated the really lenient professor. [neutral prime]
- b. They appreciated the professor who was really lenient. [*who* prime]
- c. They appreciated the professor that was really lenient. [*that* prime]

Prime sentences like (4a) contained no relative clause so it served as the neutral baseline. Sentences like (4b) and (4c) both contained relative clauses, but (4b) contained the relative pronoun *who*, and (4c) contained the overt complementizer *that*.

## Methods

### Participants

Ninety-six self-reported English-speaking monolinguals were recruited using Prolific Academic. Each participant was paid \$5 for an approximately 25-30 minute session. We doubled the sample size in Experiment 3, because the critical comparison was between the null complementizer prime condition and the neutral baseline condition as opposed to the comparison between the null complementizer prime condition and the overt complementizer prime condition, that is, we were primarily interested in the 'pure' null complementizer priming effect as opposed to the sum of the null complementizer priming effect and the *that* complementizer priming effect.

### Materials

The target sentences were identical to those used in Experiments 1 and 2. The same 24 target sentences were paired with 24 new prime sentences like (4a), (4b) and (4c), which contained identical sets of content words across conditions.

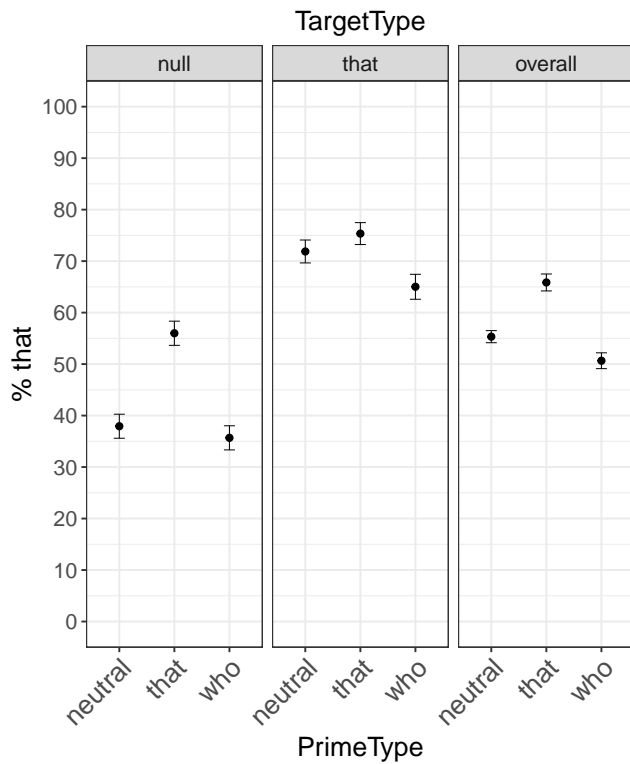
### Scoring and analysis

The scoring and analysis procedures were identical to those used in Experiments 1 and 2, except that in Experiment 3, the factor *Prime Type* was treatment-coded such that *that* primes were compared against neutral primes and null complementizer primes were compared against neutral primes.

### Results

80 out of 2304 trials (3.5%) did not meet the scoring criteria, so they were removed from further analysis.

Figure 6 shows the average rate of *that* production in each condition. Given neutral primes, speakers produced *that* 55.3% of the time. Given *who* primes, speakers produced *that* 50.7% of the time. Given *that* primes, speakers produced *that* 65.9% of the time. Thus, compared to the neutral condition, we observed 4.6% null complementizer priming and 10.6% *that* complementizer priming.

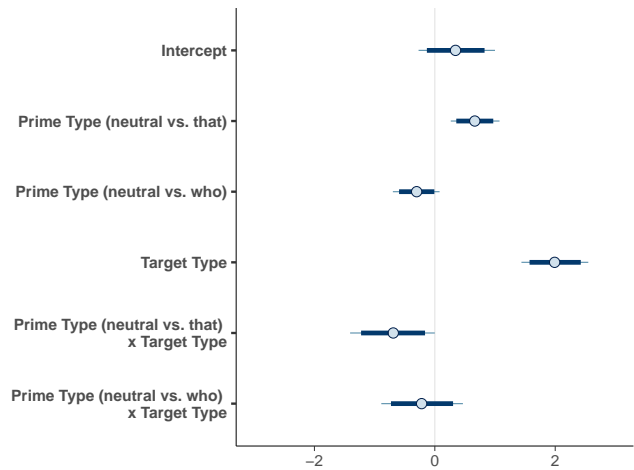


**Figure 6**

*By-condition proportions of that use in Experiment 3. Error bars represent within-subject standard errors of by-subject means.*

As can be seen in Figure 7, there was evidence for the null complementizer priming effect (the comparison between neutral and *who* primes:  $\hat{\beta} = 0.30$ , SE = 0.15, 95% CrI = [-0.59, -0.01]). There was also evidence for the *that* priming

effect (the comparison between neutral and *that* primes:  $\hat{\beta} = 0.66$ , SE = 0.15, 95% CrI = [0.36, 0.97]). Again, there was also evidence for the effect of Target Type ( $\hat{\beta} = 1.99$ , SE = 0.22, 95% CrI = [1.59, 2.42]), simply suggesting that speakers retained some memory of sentences they read. There was also evidence for the interaction between Prime Type comparing the neutral and *that* primes and Target Type ( $\hat{\beta} = -0.68$ , SE = 0.27, 95% CrI = [-1.22, -0.14]), suggesting that the *that* priming effect was primarily driven when the target sentence did not contain *that*. There was no evidence for the interaction between Prime Type comparing the neutral and *who* primes and Target Type, although numerically the null complementizer seems to be primarily driven by trials where the target sentence contained *that*. Thus, complementizer priming effects in this experiment were primarily driven by trials where target and prime structures were mismatched.



**Figure 7**

*Posterior estimates of the fixed effects with 95% (thick lines) and 99% (thin lines) credible intervals in Experiment 3.*

We conducted a post hoc analysis to see if priming effects were larger when target and prime structures were mismatched, using nested models. When target sentences contained *that*, there was evidence for the null complementizer priming effect ( $\hat{\beta} = -0.41$ , SE = 0.19, 95% CrI = [-0.78, -0.03]), but the evidence for the *that* priming effect was weak ( $\hat{\beta} = 0.31$ , SE = 0.20, 95% CrI = [-0.08, 0.72]). In contrast, when target sentences contained no *that*, there was no evidence for the null complementizer priming effect ( $\hat{\beta} = -0.19$ , SE = 0.19, 95% CrI = [-0.57, 0.18]), but there was clear evidence for the *that* priming effect ( $\hat{\beta} = 0.98$ , SE = 0.19, 95% CrI = [0.61, 1.34]). Thus, the post hoc analysis suggests that complementizer priming effects were primarily driven by trials where target and prime sentences contained mismatching complementizer structures.

## Discussion

Experiment 3 suggests that relative clauses with *who* primes the null complementizer as it occurs in VP-complement clauses, confirming the prediction of the accounts positing the null complementizer. However, this effect was relatively weak in the overall analysis (a 4.6% effect).

The post hoc analysis suggests that the null complementizer priming effect was primarily driven by the trials where target sentences contained *that*. More generally, complementizer priming effects were primarily driven by the trials where target and prime sentences contained (by hypothesis) mismatching complementizer structures. One possible reason for this is that the complementizer priming effects are subject to the *inverse preference* effect (Bock & Griffin, 2000; V. S. Ferreira, 2003; Jaeger & Snider, 2008; Reitter, Keller, & Moore, 2011). Previous studies have found that structural priming is generally stronger when a prime structure is *surprising* in a given context, and this trend seems to hold for complementizer priming specifically (V. S. Ferreira, 2003; Jaeger & Snider, 2008). What counts as *surprising* is not completely specified, but in the current context, the null complementizer or the *that* complementizer may be less surprising given an immediately preceding target sentence containing the same type of complementizer than given an immediately preceding target sentence containing the different type of complementizer. Thus, we suspect that the inverse frequency effect caused complementizer priming effects to be weaker when target and prime structures match.

## Experiment 4

In Experiment 4 we aimed to conceptually replicate the null complementizer priming effect, which was relatively small. Although a 4.6% effect is comparable in size to some of the previously observed structural priming effects in the literature (e.g., Bock and Loebell 1990 reported a 5% passive priming effect compared to the baseline in Experiment 2), we deem this replication important given that the null complementizer priming from relative clauses with *who*, if it is indeed observed, provides an important piece of evidence for the accounts assuming the null complementizer.

Because the post hoc analysis in Experiment 3 revealed that *who* relative clauses primed the null complementizer primarily when target sentences contained *that*, in Experiment 4 target sentences always contained *that*. We speculated that when target sentences contained *that*, the null complementizer is more primable due to the inverse preference effect. The design was otherwise identical to Experiment 3.

## Methods

### Participants

Seventy-two self-reported English-speaking monolinguals were recruited using Prolific Academic. Each partici-

pant was paid \$5 for an approximately 25-30 minute session. Seven participants were replaced because they did not correctly understand the task or because their audio recordings were unintelligible.

### Materials

Materials were the same as in Experiment 3, except that target sentences in Experiment 4 always contained *that*.

### Scoring and analysis

The scoring criteria and analysis procedures were identical to those in Experiment 3, except that the factor *Target Type* is no longer relevant in the analysis of Experiment 4.

## Results

95 out of 1728 trials (5.5%) did not meet the scoring criteria, so they were removed from further analysis.

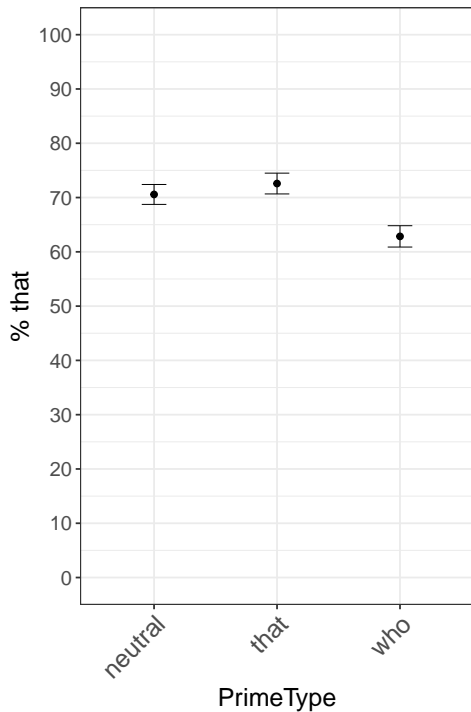
Figure 8 shows the average rate of *that* production across conditions. Given neutral primes, speakers produced *that* 70.5% of the time. Given *who* primes, they produced *that* 62.8% of the time. Given *that* primes, they produced *that* 72.6% of the time. Thus, we observed a 7.8% null complementizer priming effect and a 2.1% *that* priming effect.

As can be seen in Figure 9, there was clearer evidence for the null complementizer priming effect ( $\hat{\beta} = -0.54$ ,  $SE = 0.20$ , 95% CrI = [-0.90, -0.14]). Consistent with the results of the post hoc analysis in Experiment 3, there was no evidence for the *that* priming effect ( $\hat{\beta} = 0.14$ ,  $SE = 0.20$ , 95% CrI = [-0.26, 0.56]).

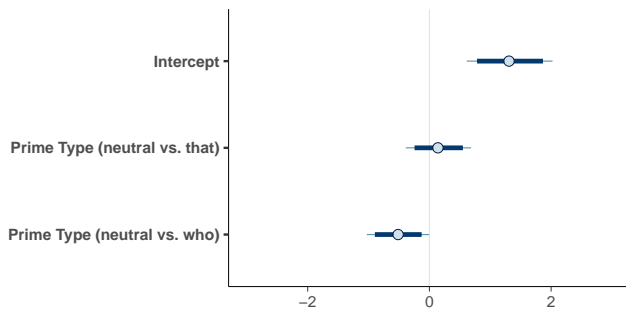
## Discussion

In Experiment 4, we observed that relative clauses with *who* primed the null complementizer as it occurs in a VP-complement clause. This replicates the results of Experiment 3, suggesting that relative clauses with *who* indeed contain the null complementizer structure.

One potential concern is that in Experiments 3 and 4, speakers might have been more likely to omit *that* when prime sentences were more complex. This is a concern because (a) neutral primes were systematically simpler than other primes, and (b) all target sentences contained *that* in Experiment 4 (unlike Experiment 3), so if speakers are more likely to forget the original sentence after reading more complex prime sentences, they may be more likely to omit *that* in non-neutral conditions, and (c) in Experiment 3, there was no clear evidence for the null complementizer priming when the target sentences did not contain *that*. Experiment 5 assesses this possibility.

**Figure 8**

By-condition proportions of *that* use in Experiment 4. Error bars represent within-subject standard errors of by-subject means.

**Figure 9**

Posterior estimates of the fixed effects with 95% (thick lines) and 99% (thin lines) credible intervals in Experiment 4.

### Experiment 5

Experiment 5 evaluated the possibility that the complexity confound was driving the effect found in Experiments 3 and 4. To this end, Experiment 5 included an additional condition where the neutral sentences included a prepositional phrase that made them more complex than the neutral sentences used in Experiments 3 and 4.

- (5) a. They appreciated the really lenient professor. [short neutral prime]
- b. They appreciated the really lenient professor in their class. [long neutral prime]
- c. They appreciated the professor that was really lenient. [*that* prime]
- d. They appreciated the professor who was really lenient. [*who* prime]

In Experiment 5, prime sentences like (5b) (long neutral primes) were added, but all other types of prime sentences (5a, 5c, and 5d) remained the same as in Experiments 3-4. If prime sentence complexity confound was driving the null complementizer priming-like pattern in Experiments 3-4, we expect that complex neutral primes should behave similarly to *who* primes.

### Methods

#### Participants

One hundred self-reported English-speaking monolingual were recruited using Prolific Academic. Each participant was paid \$5 for an approximately 25-30 minute session. Three participants were replaced because they did not follow instructions or their audio recordings were unintelligible.

#### Materials

The target sentences were identical to those in Experiments 1-4. Prime sentences were the same as in Experiments 3 and 4 except the ones used in the long neutral condition, where a sentence-final prepositional phrase was added, as in 5b.

#### Scoring and analysis

Scoring criteria and analysis procedures were identical to those in Experiments 1-4, except that Prime Type, which is a four-level factor in Experiment 5, was Helmert-coded such that long neutral primes were compared against short neutral primes (the relevant row of the contrast matrix: -1/2, 1/2, 0, 0), the average of long and short neutral primes were compared against *that* primes (-1/3, -1/3, 2/3, 0), and the average of these three conditions were compared against *who* primes (-1/4, -1/4, -1/4, 3/4).

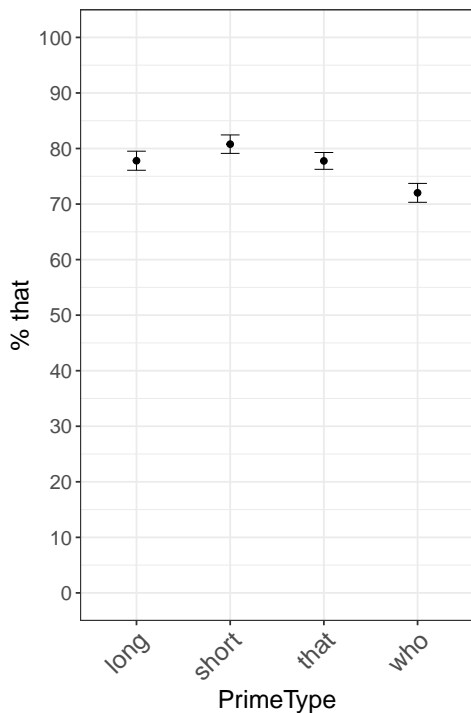
### Results

117 out of 2400 trials (4.8%) did not meet the scoring criteria so they were excluded from further analysis.

Figure 10 shows the average rate of *that* production in each condition. Given short neutral primes, speakers produced *that* 80.8% of the time. Given long neutral primes, they produced *that* 77.8% of the time. Given *that* primes, they produced *that* 77.8% of the time. Given *who* primes, they

produced *that* 72.0% of the time. Thus, given *who* primes, speakers were 5.8%, 8.8%, and 5.8% less likely to produce *that* than given short neutral primes, long neutral primes, and *that* primes, respectively.

As can be seen in Figure 11, there was evidence that *who* primes primed the null complementizer compared to the average of the other prime sentence types ( $\hat{\beta} = -0.58$ ,  $SE = 0.17$ , 95% CrI = [-0.92, -0.24]). There was no evidence that short and long neutral conditions differed ( $\hat{\beta} = 0.15$ ,  $SE = 0.25$ , 95% CrI = [-0.35, 0.63]), and no evidence that the average of the short and long neutral prime conditions differed from the *that* prime condition ( $\hat{\beta} = -0.13$ ,  $SE = 0.19$ , 95% CrI = [-0.50, 0.24]).

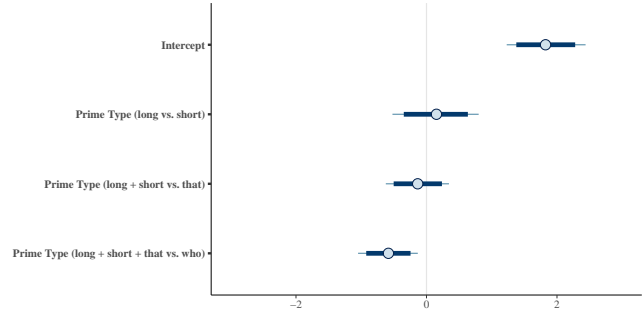


**Figure 10**

*By-condition proportions of that use in Experiment 5. Error bars represent within-subject standard errors of by-subject means.*

## Discussion

In Experiment 5, we evaluated if the null complementizer priming-like pattern in Experiment 4 was due to *who* primes being more complex than the neutral prime sentences. We found that making neutral prime sentences more complex (by adding a prepositional phrase) has a negligible effect on the rate of *that* production. There was no evidence that the 3% difference between the long and short neutral conditions was statistically meaningful, suggesting it is likely driven



**Figure 11**

*Posterior estimates of the fixed effects with 95% (thick lines) and 99% (thin lines) credible intervals in Experiment 5.*

by a small subset of participants or items. We also replicated the null complementizer priming from *who* primes. In addition, like in Experiment 4, we again failed to observe the *that* priming, suggesting that complementizer priming is observed primarily when target and prime sentences have different complementizer structures. Note that long neutral primes were arguably more complex conceptually than *who* primes because long neutral primes involved one additional concept. Long neutral primes were also phonologically longer than *who* primes, both in terms of the number of words and number of syllables. Thus, given the current results, the null complementizer priming effects found in Experiments 3-5 are unlikely to be wholly attributable to the prime complexity differences across conditions resulting in differential rates of *that*-forgetting.

## Experiment 6

Experiment 6 tested the final prediction of the accounts positing the null complementizer. Some grammatical theories in the generative tradition (e.g., Kayne 1991; Larson 1985) posit that the embedded interrogative clauses with *whether* is introduced by the null complementizer, but the embedded interrogative clauses with *if* is not. This is because *whether* is a wh-phrase and *if* is a complementizer. Under this analysis, *whether* clauses, like in *I wonder whether the cat loves the boy*, could prime the null complementizer, at least under the assumption that the null complementizer in interrogative and non-interrogative clauses are shared. We tested this prediction using prime sentences like the following.

- (6)
  - a. The scientist is unsure about the accuracy of the theory. [neutral prime]
  - b. The scientist is unsure whether the theory is accurate. [*whether* prime]
  - c. The scientist is unsure if the theory is accurate. [*if* prime]

Prime sentences like (6a) contained no embedded clause and hence served as a baseline. They contained a similar set of content words as in prime sentences in the other conditions. Prime sentences like (6b) and (6c) contained interrogative embedded clauses with either *whether* or *if*, with no other differences.

## Methods

### Participants

Seventy-two self-reported English-speaking monolinguals were recruited using Prolific Academic. Each participant was paid \$5 for an approximately 25-30 minute session. Five participants were replaced because they did not follow instructions or their recordings were unintelligible.

### Materials

The same target sentences as in Experiments 1-5 were paired with new prime sentences like (6a), (6b) and (6c). Because based on Experiment 3 the null complementizer priming may be stronger when target sentences contained *that* (possibly due to the inverse frequency effect), all target sentences contained *that* in Experiment 6, like in Experiments 4-5.

### Scoring and analysis

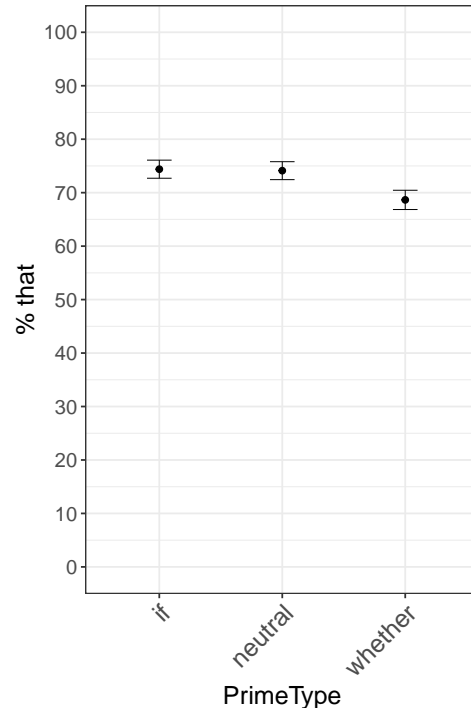
Scoring and analysis were the same as in Experiments 1-5, except that Prime Type was Helmert-coded such that *whether* primes were compared against the average of neutral and *if* primes, and *if* primes were compared against neutral primes. This coding scheme reflects the prediction that *whether* primes, but not *if* primes, prime the null complementizer.

## Results

92 out of 1728 trials (5.3%) did not meet the scoring criteria so they were excluded from further analysis.

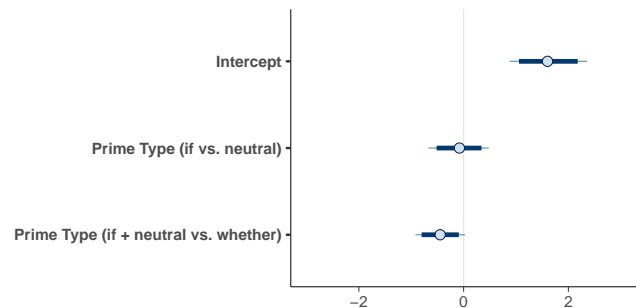
Figure 12 shows the average rate of *that* production in each condition. Given neutral primes, speakers produced *that* 74.1% of the time. Given *if* primes, they produced *that* 74.4% of the time. Given *whether* primes, they produced *that* 68.6% of the time. Thus, we observed a 5.5% priming effect.

As can be seen in Figure 13, there was evidence that *whether* sentences primed the null complementizer, compared to the average of the neutral and *if* sentences ( $\hat{\beta} = -0.45$ ,  $SE = 0.18$ , 95% CrI =  $[-0.81, -0.09]$ ). There was no evidence that *if* sentences and neutral sentences differently affected the rate of *that* production ( $\hat{\beta} = -0.08$ ,  $SE = 0.22$ , 95% CrI =  $[-0.51, 0.34]$ ).



**Figure 12**

By-condition proportions of *that* use in Experiment 6. Error bars represent within-subject standard errors of by-subject means.



**Figure 13**

Posterior estimates of the fixed effects with 95% (thick lines) and 99% (thin lines) credible intervals in Experiment 6.

## Discussion

Experiment 6 provided evidence for the final prediction of the accounts positing the null complementizer combined with some theories of grammar that assign different syntactic categories to *if* and *whether* despite very similar function they play.

Note that the issue of complexity confound in Experiment 4 does not apply to Experiment 6, because *if* and *whether*

sentences were identical except for whether they contained *if* or *whether* and yet the null complementizer priming was found. Thus, the contrast between *if* and *whether* sentences is unlikely to be explained by factors like prime sentence complexity.

### LLM evaluation

Finally, we evaluate transformer-based LLMs’ ability to capture the priming effects found in the current study. We believe this exercise is useful for two reasons. First, we can assess whether speakers’ syntactic representations go beyond observable distributional cues that LLMs likely pick up. This is relevant to the current broad theoretical question of whether our grammatical knowledge involves abstraction that goes beyond what can be gleaned purely from sensory data. Second, we can assess to what extent LLMs are good cognitive models of human linguistic performance (see Bowers et al. 2023; Huang et al. 2024; Schrimpf et al. 2021 for various perspectives).

The null complementizer is not accessible to large language models because null elements are by definition not encoded in the text inputs unless explicitly coded by humans. Also, the null complementizer priming configurations tested in the current set of studies are likely rare in naturalistic corpora. Thus, the null complementizer priming provides a strong test for LLMs as cognitive models (cf. Bowers et al. (2023)): can LLMs learn abstractions over unobservable elements with sparse data?

How can we evaluate whether and to what extent LLMs show structural priming? Sinclair, Jumelet, Zuidema, and Fernández (2022) defined the following metric:

- Priming effect =  $\log(P(T_x|P_x)) - \log(P(T_x|P_y))$

where  $T_x$  and  $T_y$  are target sentences with the structure  $x$  and  $y$ , respectively, and  $P_x$  and  $P_y$  are prime sentences with the structure  $x$  and  $y$ , respectively. Using this metric, Sinclair et al. (2022) examined three types of transformer-based LLMs, GPT-2 (Radford et al., 2019), GPT-neo (Black, Gao, Wang, Leahy, & Biderman, 2021), and DialoGPT (Zhang et al., 2019), each with different model sizes. They found that LLMs can capture basic priming effects involving passive and dative alternations, as well as cumulative priming effects (Jaeger & Snider, 2008; Kaschak & Borreggine, 2008) and lexical boost (Pickering & Branigan, 1998), although the LLMs also showed various patterns that were not human-like according to previous empirical reports.

We first used the metric introduced by Sinclair et al. (2022) to assess whether transformer-based LLMs show complementizer-priming observed in the current set of experiments. The same set of stimuli used in the current experiments were used for LLM evaluations, GPT-2, GPT-neo, and DialoGPT with three different model sizes each were assessed (Sinclair et al. 2022 additionally assessed two more

variants of GPT-2, but we omit these here to conserve space because they did not show diverging results from other models). We also do not report the results of the evaluation using the long neutral prime sentences in Experiment 5, which did not differ from the results of the evaluation using short neutral primes.

The results are summarized in Figure 14. As can be seen, using this metric, all models were able to capture the basic null complementizer priming. However, all models showed a larger priming effect within than across constructions (the top left panel, contrary to the results of Experiment 1). All models were capable of correctly distinguishing demonstrative and complementizer primes (the top right panel, in line with the results of Experiment 2). However, all models except the small DialoGPT showed the overt complementizer priming from relative clauses with *who* (the bottom left panel, compared to the neutral baseline, contrary to the results of Experiments 3 and 4), which is the opposite of the human priming pattern. Even the small DialoGPT failed to show the null complementizer priming from relative clauses with *who*. Finally, all models except GPT-neo 1.3B and GPT-neo 2.7B showed the null complementizer priming from *whether* and *if* compared to the neutral baseline. But if anything, the null complementizer priming was stronger from *if* clauses than from *whether* clauses (the bottom right panel, contrary to the results of Experiment 5). In fact, the null complementizer priming was systematically stronger from *if* clauses, unlike humans. GPT-neo 1.3B showed the overt complementizer priming from both *if* and *whether* primes compared to the neutral baseline, unlike humans.

Next, we used *surprisal* of the overt complementizer *that* in target sentences, given different types of prime sentences as context. The complementizer priming effect under this approach is defined as the difference in surprisal of *that* given two different prime sentences. That is, it would be manifested as a higher surprisal of *that* in the target sentence given a prime sentence with *that* than a minimally different prime sentence without *that*. Note that this approach works for complementizer priming because the structural alternation is localizable to one word. The results are shown in Figure 15. Perhaps not surprisingly, the overall pattern looks quite similar to the evaluations based on the metric by Sinclair et al. (2022). All models investigated here showed within- and cross-constructional complementizer priming, but the magnitude of the effect was smaller across than within constructions (the top left panel in Figure 15, contrary to the results of Experiment 1), possibly except the DialoGPT-large (the same numerical pattern was still there, but it was less clear in this model). All models investigated here also showed priming that is not reducible to the mere presence or absence of the phonological form, *that* (the top right panel in Figure 15, in line with the results of Experiment 2). However, virtually all LLMs investigated here (possibly except the DialoGPT-



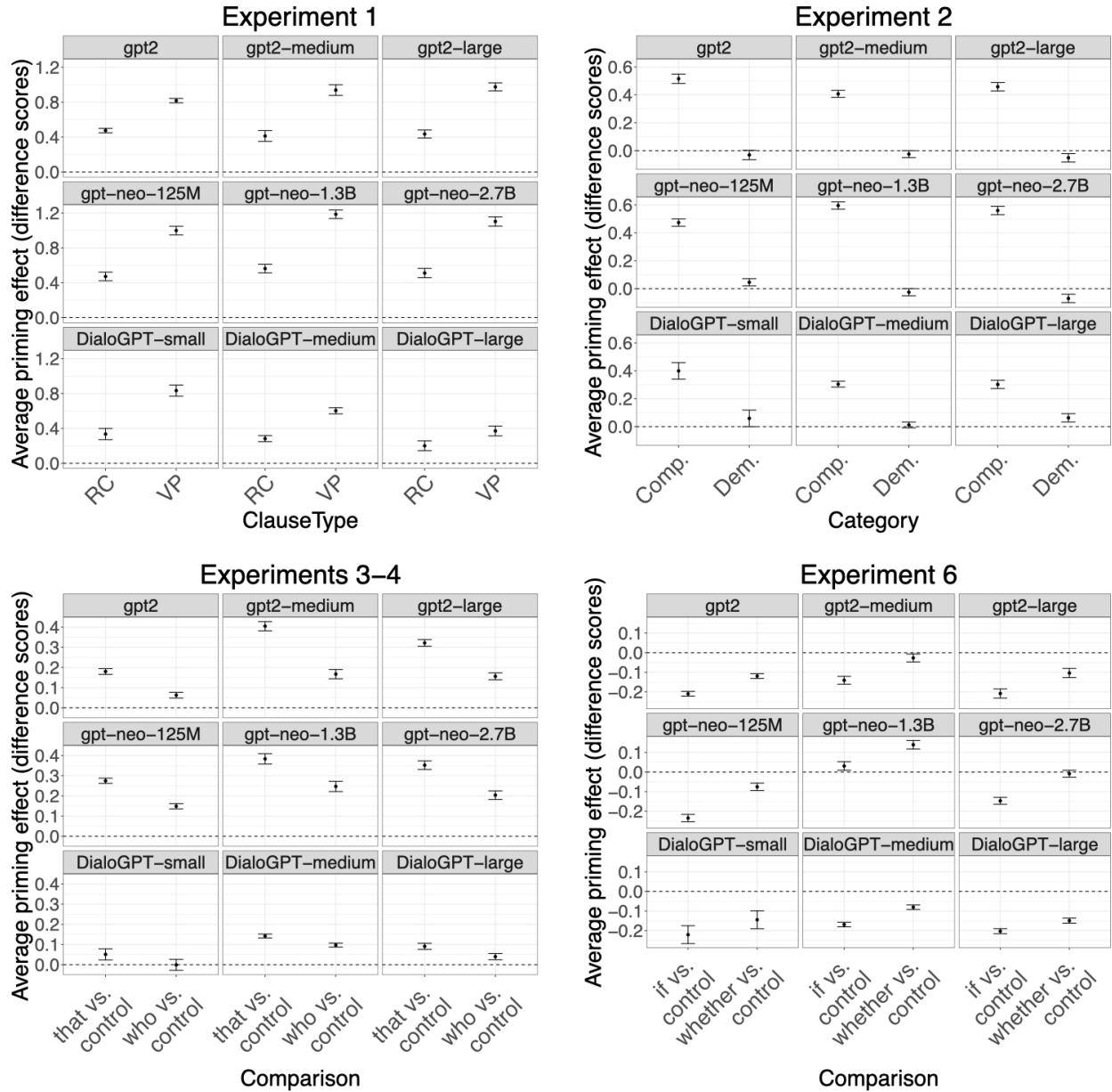


Figure 14

The priming effects in LLMs based on the metric proposed by Sinclair et al. (2022). Note that the y-axes all indicate difference scores (i.e., the priming effect), and for Experiments 3-5 where the factor Prime Type had three-level, non-neutral primes were each compared against neutral primes. Dotted horizontal line indicates 0 (no effect). Error bars represent standard errors.

small, which showed no priming effect in Experiments 3-4 whatsoever) showed that *who* primes the overt, not null, complementizer unlike humans (the bottom left panel in Figure 15, contrary to the results of Experiments 3 and 4). Also, all models except the DialoGPT-small showed that *whether*

primes *that* compared to *if* primes (the bottom right panel in Figure 15, contrary to the results of Experiments 6). The only exception - DialoGPT-small - also failed to show the human-like pattern: it showed that *if* and *whether* both primed the null complementizer, which again is not consistent with hu-



man priming effects.

Overall, the current evaluation suggests that the parallelism between transformer-based LLMs and humans is limited. Although LLMs exhibited the basic complementizer priming effect (Experiments 1 and 2) and captured categorical distinction of homophonous *thats* (Experiment 2), they diverged from humans in at least two important ways. First, LLMs were more sensitive to constructional differences than humans, showing a larger priming effect within than across constructions (contra Experiment 1). Second, they showed no evidence that they captured the generalization across a diverse sets of embedded clauses lacking *that*, including relative clauses with *who* and complement clauses with *whether* (contra Experiments 3-6). There is no obvious systematic relation between model size and the behaviors of LLMs, suggesting that simply scaling up (or down) the models is unlikely to make the alignment better. In sum, the kind of abstraction transformer-based LLMs make about (embedded) clausal structures seem distinct from the kind of abstractions humans make.

The current results of course do not suggest that LLMs can never be human-like in terms of how they encode grammatical structures of embedded clauses. It may be that the metrics we use to evaluate LLMs (here, the definition of the priming effect provided by Sinclair et al. 2022 and the surprisal of *that* in target sentences) may not be adequate or that LLMs become more human-like over time with changes in their architectures or training regimen. Nevertheless, if LLMs were to be evaluated as cognitive models, they need to be evaluated against the phenomena that have a high chance of challenging them (Bowers et al., 2023), such as the null complementizer priming effects we report in the current study.

### General Discussion

We tested whether speakers encode the null complementizer using structural priming. Speakers were primed for the null complementizer from relative clauses to VP-complement clauses just as much as from VP-complement clauses to VP-complement clauses (Experiments 1 and 2), suggesting that speakers encode an abstract representation of complementizers above and beyond the mere absence of the overt complementizer. We also found that speakers were primed for the null complementizer (as it occurs in VP-complement clauses) from relative clauses with *who* (Experiments 3-5) compared to the neutral baselines. This is unlikely to be due to the complexity differences between the neutral primes and *who* primes (Experiment 5). Speakers were also primed for the null complementizer from interrogative complement clauses with *whether*, but not from minimally different interrogative complement clauses with *if* (Experiment 6). These results overall suggest that speakers indeed encode the null complementizer in a manner that closely aligns with grammatical analyses of embedded

clausal structures under some grammatical theories assuming null elements.

There are at least two broad theoretical implications of the current results. The first is that speakers encode abstract syntactic representations devoid of perceptual-motor content. As discussed in the introduction, whether speakers and comprehenders encode abstract representations devoid of perceptual-motor content is controversial in linguistics and psycholinguistics. However, we contend that the current set of results is difficult to explain without positing the null complementizer. In agreement with Phillips and Wagers (2007), we believe this conclusion is quite natural from the perspective of broader cognitive science. It is uncontroversial that perception and cognition routinely go beyond distal stimuli in non-linguistic domains of cognition. For instance, the visual perception of the Kanizsa triangle involves perceiving illusory contours, constructed based on our visual system's implicit knowledge. Likewise, speakers encode the null complementizer based on the grammatical knowledge of clausal structures. In our assessment, the current set of experiments provides overall strong experimental evidence that speakers indeed encode the null complementizer, a textbook case of abstract representations devoid of perceptual-motor content.

Of course, this does not mean that other types of null elements (such as traces/copies) must also be encoded in sentence production (see Bock, Loebell, and Morey 1992 for an argument against the involvement of movement transformation and F. Ferreira and Engelhardt 2006 for a discussion of null elements in production). Each type of null element must be evaluated separately. However, the current results at least suggest that a general ban against null elements can make grammatical theories less, not more, psychologically adequate. That is, *WYSIWYG* is not necessarily a virtue as far as psychological adequacy is concerned.

A related implication of the current model is that syntactic representations are not reducible to other factors such as meaning or frequency (cf. *Autonomy of Syntax*, see Adger 2018; Chomsky 1957; Newmeyer 2017; Partee 2014). The reason is that the complementizer priming effects we observed here seem irreducibly syntactic. We are aware of no phonological, semantic, or pragmatic factors that are selectively shared across sentences that primed the null complementizer (relative clauses with *who*, relative clauses without relative pronouns or *that*, VP-complementizer clauses with the null complementizer, and *whether* clauses), to the exclusion of sentences that didn't (relative clauses with *that*, VP-complement clauses with *that*, and *if* clauses). Of course, us not being aware of such a factor is not proof of non-existence. However, if one wants to explain away the current results in non-syntactic terms, the challenge is to specify exactly what such a non-syntactic factor is, and how it captures the current set of results (as well as past results). In the production

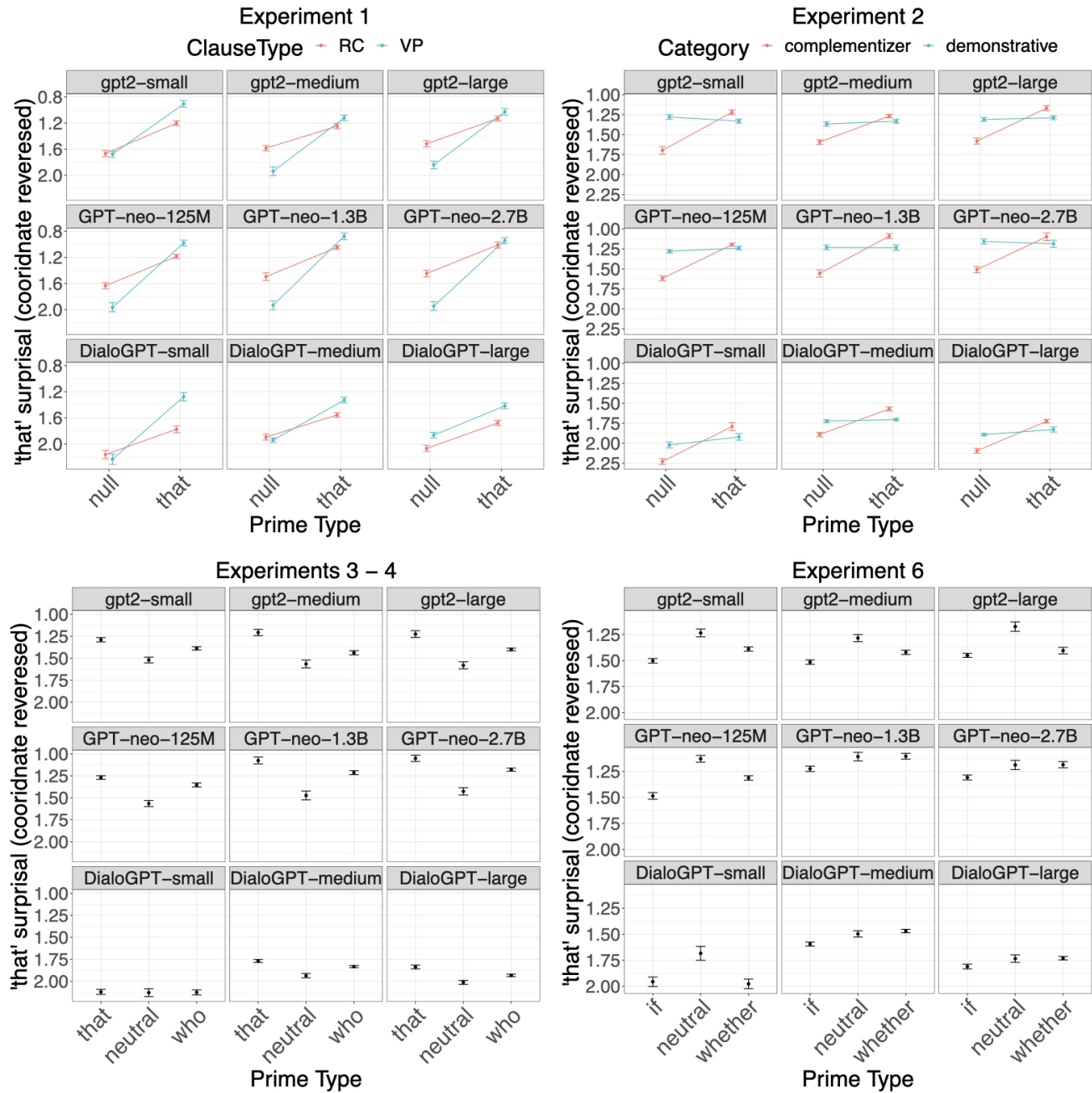


Figure 15

The surprisal of the overt complementizer in target sentences derived from various LLMs given various prime sentence types across experiments. Note that the y-axes were all inverted for easy comparison with human data presented above. Error bars represent standard errors.

literature, there is long-standing resistance against adopting syntactic representations as explanatory constructs, and there have been attempts to explain away various structural priming effects in semantic, pragmatic or prosodic terms (see Ziegler et al. 2019 for a recent example). To the extent the

current results resist non-syntactic explanations, they suggest that speakers indeed encode autonomous syntactic representations that are not reducible to other factors.

It is worth noting that a prominent account of structural priming explaining structural priming in terms of *combina-*

*torial potential* of a lexical item struggles to capture the current results. In particular, in Reitter et al. (2011) (see also Prasad, Van Schijndel, and Linzen 2019 for a related view), dative priming is thought to be the priming of a combinatory category of dative verbs. Prepositional dative licensing verbs have the combinatory category (VP / PP) / NP (informally, this means that a dative verb taking a prepositional object is a function that takes NP to the right to yield a function that takes PP to the right to yield a VP). For double object dative licensing verbs have the distinct combinatory category (VP / NP) / NP. The general idea behind this approach is that these combinatorial categories are the locus of structural priming. This approach to structural priming can capture within-construction complementizer priming. Clause-taking verbs like *think* may have two distinct combinatorial categories (specifically, VP / CP and VP / TP) and thus one could say that these combinatorial categories can be primed. However, this combinatorial category-based account faces a serious challenge as soon as across-construction complementizer priming needs to be taken into account. This is because combinatorial categories of verbs introducing VP-complement clauses and combinatorial categories of nouns that relative clauses modify are clearly distinct (VP / S vs. simply N). In addition, the prominent subclass of CCG avoids null elements (the principle of adjacency; see Steedman 2001, p. 54), and consequently must assimilate relative pronouns (*who*) and the overt complementizer *that* as it occurs in relative clauses because both *who* and *that* occur in the same distributional environment. The same holds for the contrast between *if* and *whether*. The current results are difficult to explain under the assumption that *who* and *that* on one hand, and *whether* and *if* on the other share the combinatorial categories. Therefore, CCG-based accounts of structural priming, to the extent they avoid positing the null complementizer, struggle to capture the current set of results.

It is an open question whether the model like Chang, Dell, and Bock (2006) can capture the current results. A noteworthy relevant finding in the current study is that the null complementizer priming does not seem to require *optionality* of the priming structures. The implicit learning theory of structural priming, which Chang et al. (2006) is a species of, generally predicts that priming occurs to the extent that more or less the same message can be expressed in multiple ways, i.e., when a structure is optional. However, in the current study, the null complementizer priming occurred from *whether* but not *if* clauses, and neither allows message-preserving syntactic alternations. Thus, at least *prima facie*, the implicit learning theory of priming may not be sufficient to capture structural priming in its entirety. However, this of course does not mean that implicit learning plays no role in priming; it just means that it may not be sufficient to explain various types of structural priming effects. Also, eval-

uating an implemented computational model like Chang et al. (2006) requires careful simulations, which are beyond the scope of the current paper. Because complex models like Chang et al. (2006) may not align with intuitions about how they should work, we leave this as a topic for future research.

Turning to the results of LLM evaluation, the current results suggest that the parallelism between transformer-based LLMs and humans is limited. This adds to the emerging evidence that LLMs do not adequately capture various psycholinguistic effects when tested with infrequent phenomena, at least when assessed using surprisal as a metric of evaluation (Huang et al., 2024), and it is consistent with the report by Sinclair et al. (2022), which also showed some divergence between LLMs and humans in terms of structural priming.

Simple structural priming effects do not offer a particularly strong test of LLMs as cognitive models because various types of corpora clearly contain information reflecting structural priming effects and associated modulatory effects. The basic priming effect, the cumulativeness effect, the lexical boost effect, and the inverse frequency effect can all be statistically detected in corpora (Gries, 2005; Jaeger & Snider, 2008; Reitter et al., 2011), meaning that relevant configurations occur relatively frequently. LLMs can pick up such regularity in the corpora. Certainly, capturing these basic patterns requires some abstraction over surface strings and LLMs are clearly capable of doing so, and this itself is an interesting capacity of LLMs. However, if LLMs were to be construed as a cognitive model, the mere fact that LLMs make some sort of abstraction is far from sufficient. We must evaluate what kinds of abstraction LLMs make, and whether they align with abstractions humans make. To assess the alignment, it is necessary to look beyond patterns that frequently occur in training data (see Huang et al. 2024). Null complementizer priming offers a strong test of LLMs as cognitive models in that the relevant structural configurations are likely rare in naturalistic settings, and the null complementizer is not visible in input strings and hence its learning requires a considerable degree of abstraction. As it currently stands, LLMs do not pass this strong test, suggesting that the abstractions that LLMs and humans make with respect to embedded clause structures do not align with each other. More specifically, the current results point to the conclusion that LLMs make less abstract generalizations across constructions than humans in that (a) they are more sensitive to differences across constructions and (b) they are less sensitive to abstract similarities across surface distinct constructions.

## Conclusion

Overall, the current results suggest that speakers encode the null complementizer in a manner that closely aligns with analyses under some grammatical theories positing null elements and in a manner that diverges from what LLMs learn

from distributional cues. From these results, we conclude that mental representations of sentences that speakers construct during sentence production go beyond what we hear and say.

### Appendix Stimulus list

The list of target-prime sentence pairs used in the current study. (a): target sentences (shared across all experiments). (b), (c), (d), (e): prime sentences in Experiments 1, 2, 3-5, 6, respectively. The long natural prime sentences used in Experiment 5 included prepositional phrases at the end of each neutral prime sentences. The same set of sentences were used for LLM evaluations.

- (1)
  - a. The director announced (that) the actor would be in the new movie. (Target)
  - b. The professor thought (that) the student appreciated the idea during the class / The professor appreciated the thoughts (that) the students expressed during the class.
  - c. The professor appreciated the thoughts (that) the student expressed / The professor appreciated the thoughts of (the/that) student.
  - d. They appreciated the really lenient professor (in their department). / They appreciated the professor that was really lenient / They appreciated the professor who was really lenient.
  - e. The scientist is unsure if the theory is accurate / The scientist is unsure about the accuracy of the theory / The scientist is unsure whether the theory is accurate.
- (2)
  - a. The breeder believed (that) the dog was well-trained.
  - b. Phil claimed (that) his roommate lied during the party / Phil is angry about the claim (that) his roommate made during the party.
  - c. Phil was angry about the claim (that) the rude customer made / Phil was angry about the claim by (the / that) rude customer.
  - d. Phil was angry about the loud customer (at the restaurant) / Phil was angry about the customer that was loud / Phil was angry about the customer who was loud.
  - e. The judge is not confident about the witness's reliability / The judge is not confident if the witness is reliable / The judge is not confident whether the witness is reliable.
- (3)
  - a. The student said (that) the teacher was nice.
  - b. The pilot announced (that) the flight was going to be delayed for an hour / Emily heard the announcement (that) the pilot made during the flight.
- c. Emily heard the announcement (that) the flight attendant made / Emily heard the announcement from (the / that) flight attendant.
- d. Emily didn't like the really talkative flight attendant (on the long flight) / Emily didn't like the flight attendant that was really talkative / Emily didn't like the flight attendant who was really talkative.
- e. The journalist checked the reliability of the information source / The journalist checked if the information source is reliable / The journalist checked whether the information source is reliable.
- (4)
  - a. The little girl heard (that) the bird visited her house.
  - b. The teacher argued (that) the student wrote the best paper / The teacher praised the argument (that) the student made in her paper.
  - c. The teacher praised the argument (that) the student made / The teacher praised the argument of (the / that) student.
  - d. The little girl admires the extremely intelligent woman (from her neighborhood) / The little girl admires the woman that is extremely intelligent / The little girl admires the woman who is extremely intelligent.
  - e. The mother inquired about the cost of the summer camp / The mother inquired if the summer camp costs money / The mother inquired whether the summer camp costs money.
- (5)
  - a. The dentist thought (that) people needed to floss more often.
  - b. The activist remarked (that) the politician made an inappropriate comment some time ago / The activist criticized the remark (that) the politician made some time ago.
  - c. The activist criticized the remark (that) the politician made / The activist criticized the remark of (the / that) politician.
  - d. The activist criticized the generally unreliable politician (during the rally) / The activist criticized the politician that is generally unreliable / The activist criticized the politician who is generally unreliable.
  - e. The doctor is not confident about the accuracy of his diagnosis / The doctor is not confident if his diagnosis is accurate / The doctor is not confident whether his diagnosis is accurate.

- (6) a. The nurse claimed (that) the patient had fainted.  
 b. The historian discovered (that) his theory was wrong during the trip / The historian wrote about the discovery (that) he made during the trip.  
 c. The historian wrote about the discovery (that) the anthropologist made / The historian wrote about the discovery of (the / that) anthropologist.  
 d. The journalist interviewed the extremely knowledgeable historian (about the recent discovery) / The journalist interviewed the historian that is extremely knowledgeable / The journalist interviewed the historian who is extremely knowledgeable.  
 e. The CEO checked the appropriateness of the statement / The CEO checked if the statement was appropriate / The CEO checked whether the statement was appropriate.
- (7) a. The engineer suspected (that) the software had a lot of bugs.  
 b. The scientist asserted (that) his peer got the facts wrong during the conference / The scientist was skeptical about the assertion (that) his peer made during the conference.  
 c. The scientist was skeptical about the assertion (that) the speaker made / The scientist was skeptical about the assertion by (the / that) speaker.  
 d. The scientist wasn't impressed by the clearly dishonest speaker (at the conference) / The scientist wasn't impressed by the speaker that is clearly dishonest / The scientist wasn't impressed by the speaker who is clearly dishonest.  
 e. The engineer questioned the feasibility of the new project / The engineer questioned if the new project is feasible / The engineer questioned whether the new project is feasible.
- (8) a. The chef feared (that) his cake was too sweet.  
 b. The wedding planner proposed (that) the bride's family visit the chapel / The bride liked the proposal (that) the wedding planner made.  
 c. The bride liked the proposal (that) the wedding planner made / The bride liked the proposal by (the / that) wedding planner.  
 d. The bride liked the really well-organized wedding planner (for the ceremony) / The bride liked the wedding planner that is really well-organized / The bride liked the wedding planner who is really well-organized.  
 e. The homeowner wondered about the value of his property / The homeowner wondered if his property is still valuable / The homeowner wondered whether his property is still valuable.
- (9) a. The boy guessed (that) the computer was broken.  
 b. The travel agent confirmed that the couple had a hotel reservation / The travel agent double-checked the confirmation the hotel sent out.  
 c. The travel agent double-checked the confirmation (that) the hotel sent out / The travel agent double-checked the confirmation from (the / that) hotel.  
 d. The travel agent has difficulty dealing with the extremely unpredictable customer (on the phone) / The travel agent has difficulty dealing with the customer that is extremely unpredictable / The travel agent has difficulty dealing with the customer who is extremely unpredictable.  
 e. The professor is unsure about the student's understanding of the topic / The professor is unsure if the student understood the topic / The professor is unsure whether the student understood the topic.
- (10) a. The professor argued (that) the tuition should be cheaper.  
 b. The police officer warned (that) the driver would get a ticket / The driver ignored the warning (that) the police officer issued.  
 c. The driver ignored the warning (that) the police officer issued / The driver ignored the warning from (the / that) police officer.  
 d. The driver ignored the very aggressive police officer (at the checkpoint) / The driver ignored the police officer that was very aggressive / The driver ignored the police officer who was very aggressive.  
 e. The investor questioned the company's financial stability / The investor questioned if the company is financially stable / The investor questioned whether the company is financially stable.
- (11) a. The student assumed (that) the exam was easy.  
 b. The fortune teller predicted (that) the woman would win the lottery / The woman laughed at the prediction (that) the fortune teller made.  
 c. The woman laughed at the prediction (that) the fortune teller made / The woman laughed at the prediction by (the / that) fortune teller.  
 d. The woman believed the famous fortune teller (in the city) / The woman believed the fortune teller that is famous / The woman believed the fortune teller who is famous.  
 e. The detective investigated the validity of the suspect's alibi / The detective investigated if the suspect's alibi is valid / The detective investigated whether the suspect's alibi is valid.

- (12) a. The novelist implied (that) the new character would die.  
 b. The dean reported (that) the program needs more staff / The dean received the report (that) the program chair wrote.  
 c. The dean received the report (that) the professor wrote / The dean saw the report from (the / that) professor.  
 d. The dean didn't like the difficult professor (in the math department) / The dean didn't like the professor that was difficult / The dean didn't like the professor who was difficult.  
 e. The auditor assessed the accuracy of the financial statements / The auditor assessed if the financial statements were accurate / The auditor assessed whether the financial statements were accurate.
- (13) a. The girl believed (that) the library was closed.  
 b. The investigator concluded (that) the suspect was guilty / The investigator accepted the conclusion (that) his assistant drew.  
 c. The judge trusted the detailed description (that) the witness provided / The judge trusted the detailed description by (the / that) witness.  
 d. The judge trusted the apparently honest witness (in the trial) / The judge trusted the witness that seemed honest / The judge trusted the witness who seemed honest.  
 e. The economist is unsure about the effectiveness of the new policy / The economist is unsure if the new policy is effective / The economist is unsure whether the new policy is effective.
- (14) a. The manager implied (that) no one was being laid off.  
 b. The student guessed (that) the teacher would make the exam difficult / The teacher was surprised by the guess (that) the student made.  
 c. The teacher was surprised by the proposal (that) the student described / The teacher was surprised by the proposal by (the / that) student.  
 d. The student was amazed by the charismatic teacher (in her class) / The student was amazed by the teacher that was charismatic / The student was amazed by the teacher who was charismatic.  
 e. The counselor assessed the emotional stability of the child / The counselor assessed if the child is emotionally stable / The counselor assessed whether the child is emotionally stable.
- (15) a. The author said (that) the book sold well.  
 b. The broker advised (that) low-risk investors should sell their stock. / The client didn't listen to the advice (that) the broker offered.  
 c. The client didn't listen to the advice (that) the broker offered / The client didn't listen to the advice from (the / that) broker.  
 d. The client disliked the arrogant broker (at the firm) / The client disliked the broker that was arrogant / The client disliked the broker who was arrogant.  
 e. The researcher questioned the robustness of the finding / The researcher questioned if the finding is robust / The researcher questioned whether the finding is robust.
- (16) a. The fireman thought (that) the boy had already been rescued.  
 b. The company stated (that) its employees are allowed to work remotely / The company endorsed the statement (that) the employee made.  
 c. The company endorsed the statement (that) the employee made / The company endorsed the statement by (the / that) employee.  
 d. The company hired the extremely creative engineer (for the new project) / The company hired the engineer that is extremely creative / The company hired the engineer who is extremely creative.  
 e. The analyst evaluated the impact of the marketing campaign / The analyst evaluated if the marketing campaign made a difference / The analyst evaluated whether the marketing campaign made a difference.
- (17) a. The doctor heard (that) the patient was angry.  
 b. John irrationally believes (that) his sister is a spy / John shares the irrational belief (that) her sister holds.  
 c. John shares the irrational belief (that) the girl holds / John shares the irrational belief of (the / that) girl.  
 d. John was astonished by the very eloquent girl (at the debate) / John was astonished by the girl that was very eloquent / John was astonished by the girl who was very eloquent.  
 e. The coach is unsure about the effectiveness of the new training program / The coach is unsure if the new training program is effective / The coach is unsure whether the new training program is effective.
- (18) a. The historian claimed (that) the king was well respected.

- b. The author acknowledged (that) he didn't proofread his draft / The author proofread the acknowledgment (that) he wrote.
- c. The editor proofread the acknowledgment (that) the author wrote / The editor proofread the acknowledgment by (the / that) author.
- d. The author called the usually supportive editor (in the evening) / The author called the editor that is usually supportive / The author called the editor who is usually supportive.
- e. The lawyer is not confident about the validity of the argument / The lawyer is not confident if the argument is valid / The lawyer is not confident whether the argument is valid.
- (19) a. The woman guessed (that) her husband wanted to quit his job.
- b. Alison suggested (that) her brother get a haircut / Alison liked the suggestion (that) her brother made.
- c. Alison followed the suggestion (that) the lawyer made / Alison followed the suggestion by (the / that) lawyer.
- d. Alison trusts the extraordinarily empathetic lawyer (with her case) / Alison trusts the lawyer that is extraordinarily empathetic / Alison trusts the lawyer who is extraordinarily empathetic.
- e. The veterinarian assessed the general health of the dog / The veterinarian assessed if the dog was generally healthy / The veterinarian assessed whether the dog was generally healthy.
- (20) a. The farmer suspected (that) the florist was already married.
- b. The psychologist felt (that) her client was doing just fine after the session / The psychologist understood the feelings (that) her client shared with her.
- c. The psychologist understood (that) the feelings the client expressed / The psychologist understood the feelings of (the / that) client.
- d. The psychologist finally understood the extremely shy client (after several sessions) / The psychologist finally understood the client that is extremely shy / The psychologist finally understood the client who is extremely shy.
- e. The activist asked about the politician's view on environmental issues / The activist asked if the politician cares about environmental issues / The activist asked whether the politician cares about environmental issues.
- (21) a. The author feared (that) his new book would be criticized harshly.
- b. The librarian noted (that) the visitor had an overdue book / The visitor received the note (that) the librarian wrote.
- c. The visitor received the note (that) the receptionist wrote / The visitor received the note from (the / that) receptionist.
- d. The visitor talked to the really polite receptionist (at the front desk) / The visitor talked to the receptionist that was really polite / The visitor talked to the receptionist who was really polite.
- e. The traveler asked about the fitness center at the hotel / The traveler asked if the hotel has a fitness center / The traveler asked whether the hotel has a fitness center.
- (22) a. The union leader argued (that) the employees should be paid more.
- b. The doctor observed (that) the patient could not stand up straight / The doctor appreciated the observation (that) the nurse made about the patient.
- c. The doctor appreciated the observation (that) the nurse made / The doctor appreciated the observation by (the / that) nurse.
- d. The doctor appreciated the very detail-oriented assistant (during the surgery) / The doctor appreciated the assistant that is very detail-oriented / The doctor appreciated the assistant who is very detail-oriented.
- e. The sheriff wondered about the safety of the neighborhood / The sheriff wondered if the neighborhood is safe / The sheriff wondered whether the neighborhood is safe.
- (23) a. The president announced (that) the company would donate to the charity.
- b. The therapist noticed (that) his patient canceled the appointment yesterday / The patient received the billing notice (that) his therapist issued.
- c. The patient received the billing notice (that) the therapist issued / The patient received the billing notice from (the / that) therapist.
- d. The patient hated the really condescending therapist (in the clinic) / The patient hated the therapist that was really condescending / The patient hated the therapist who was really condescending.
- e. The architect wondered about the stability of the building's foundation / The architect wondered if the building's foundation is stable / The architect wondered whether the building's foundation is stable.
- (24) a. The midwife assumed (that) the baby was a boy.

- b. The gym instructor explained (that) the equipment must be cleaned after each use / The customer didn't really understand the explanation (that) the gym instructor provided.
- c. The customer didn't understand the explanation (that) the gym instructor provided / The customer didn't understand the explanation by (the / that) gym instructor.
- d. Bill had difficulty hearing the really quiet instructor (in the lecture hall) / Bill had difficulty hearing the instructor that was really quiet / Bill had difficulty hearing the instructor who was really quiet.
- e. Alison wondered about the effectiveness of the treatment / Alison wondered if the new treatment is effective / Alison wondered whether the new treatment is effective.

#### References

- Abney, S. P. (1987). *The english noun phrase in its sentential aspect* (Unpublished doctoral dissertation). Massachusetts Institute of Technology.
- Adger, D. (2018). Autonomy of syntax. In N. Hornstein, H. Lasnik, P. Patel-Grosz, & C. Yang (Eds.), *The autonomy of syntax* (Vol. 129, p. 153-176). Walter de Gruyter.
- Aelbrecht, L. (2010). *The syntactic licensing of ellipsis*. John Benjamins.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278.
- Bhatt, R. (2002). The raising analysis of relative clauses: Evidence from adjectival modification. *Natural language semantics*, 10(1), 43–90.
- Bhatt, R., & Pancheva, R. (2006). Implicit arguments. *The Blackwell companion to syntax*, 558–588.
- Black, S., Gao, L., Wang, P., Leahy, C., & Biderman, S. (2021). Gpt-neo: Large scale autoregressive language modeling with mesh-tensorflow.
- Bock, K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, 18(3), 355–387.
- Bock, K. (1989). Closed-class immanence in sentence production. *Cognition*, 31(2), 163–186.
- Bock, K., Dell, G. S., Chang, F., & Onishi, K. H. (2007). Persistent structural priming from language comprehension to language production. *Cognition*, 104(3), 437–458.
- Bock, K., & Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology: General*, 129(2), 177.
- Bock, K., & Loebell, H. (1990). Framing sentences. *Cognition*, 35(1), 1–39.
- Bock, K., Loebell, H., & Morey, R. (1992). From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological Review*, 99(1), 150.
- Bošković, Ž., & Lasnik, H. (2003). On the distribution of null complementizers. *Linguistic Inquiry*, 34(4), 527–546.
- Bowers, J. S., Malhotra, G., Adolphi, F., Dujmović, M., Montero, M. L., Biscione, V., ... Heaton, R. F. (2023). On the importance of severely testing deep learning models of cognition. *Cognitive Systems Research*, 82, 101158.
- Branigan, H. P., & Pickering, M. J. (2017). An experimental approach to linguistic representation. *Behavioral and Brain Sciences*, 40.
- Bresnan, J., Asudeh, A., Toivonen, I., & Wechsler, S. (2015). *Lexical-functional syntax*. Malden, MA: John Wiley & Sons.
- Browning, M. (1987). *Null operator constructions* (Unpublished doctoral dissertation). Massachusetts Institute of Technology.
- Bürkner, P.-C. (2021). Bayesian item response modeling in R with brms and Stan. *Journal of Statistical Software*, 100(5), 1–54. doi: M10.18637/jss.v100.i05
- Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, 113(2), 234.
- Chomsky, N. (1957). *Syntactic structure*. The Hague: Mouton.
- Chomsky, N. (1981). *Lectures on government and binding: The pisa lectures*. Dordrecht: Foris Publications.
- Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.
- Chomsky, N., & Lasnik, H. (1977). Filters and control. *Linguistic inquiry*, 8(3), 425–504.
- Contreras, H. (1984). A note on parasitic gaps. *Linguistic Inquiry*, 15(4), 698–701.
- Culicover, P. W., & Jackendoff, R. (2006). The simpler syntax hypothesis. *Trends in cognitive sciences*, 10(9), 413–418.
- De Marneffe, M.-C., & Nivre, J. (2019). Dependency grammar. *Annual Review of Linguistics*, 5, 197–218.
- Ferreira, F., & Engelhardt, P. (2006). In M. Traxler & M. Gernsbacher (Eds.), *Handbook of psycholinguistics* (pp. 61–91).
- Ferreira, V. S. (2003). The persistence of optional complementizer production: Why saying “that” is not saying “that” at all. *Journal of Memory and Language*, 48(2), 379–398.
- Gazdar, G. (1981). Unbounded dependencies and coordinate structure. *Linguistic Inquiry*, 12, 155–184.
- Gibson, E., & Hickok, G. (1993). Sentence processing with empty categories. *Language and Cognitive Processes*,



- 8(2), 147–161.
- Ginzburg, J., & Sag, I. (2000). *Interrogative investigations*. Stanford: CSLI publications.
- Goldberg, A., & Suttle, L. (2010). Construction grammar. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(4), 468–477.
- Gorrell, P. (1993). Evaluating the direct association hypothesis: A reply to pickering and barry (1991). *Language and Cognitive Processes*, 8(2), 129–146.
- Gries, S. T. (2005). Syntactic priming: A corpus-based approach. *Journal of psycholinguistic research*, 34, 365–399.
- Haik, I. (1985). *The syntax of operators* (Unpublished doctoral dissertation). Massachusetts Institute of Technology.
- Hornstein, N. (1999). Movement and control. *Linguistic Inquiry*, 30(1), 69–96.
- Huang, K.-J., Arehalli, S., Kugemoto, M., Muxica, C., Prasad, G., Dillon, B., & Linzen, T. (2024). Large-scale benchmark yields no evidence that language model surprisal explains syntactic disambiguation difficulty. *Journal of Memory and Language*, 137, 104510.
- Jaeger, T. F., & Snider, N. (2008). Implicit learning and syntactic persistence: Surprisal and cumulativity. In *Proceedings of the 30th annual conference of the cognitive science society* (Vol. 827812).
- Kanizsa, G., et al. (1955). Margini quasi-percettivi in campi con stimolazione omogenea. *Rivista di psicologia*, 49(1), 7–30.
- Kaplan, R. M., & Bresnan, J. (1981). *Lexical-functional grammar: A formal system for grammatical representation*. Massachusetts Institute Of Technology, Center For Cognitive Science.
- Kaschak, M. P., & Borreggine, K. L. (2008). Is long-term structural priming affected by patterns of experience with individual verbs? *Journal of Memory and Language*, 58(3), 862–878.
- Kayne, R. S. (1991). Romance clitics, verb movement, and pro. *Linguistic inquiry*, 22(4), 647–686.
- Kishimoto, H. (2006). On the existence of null complementizers in syntax. *Linguistic Inquiry*, 37(2), 339–345.
- Kratzer, A. (1996). Severing the external argument from its verb. In J. Rooryck & L. Zarin (Eds.), *Phrase structure and the lexicon* (pp. 109–137). Dordrecht, the Netherlands: Springer.
- Larson, R. K. (1985). On the syntax of disjunction scope. *Natural Language & Linguistic Theory*, 3(2), 217–264.
- Levelt, W. J., & Kelter, S. (1982). Surface form and memory in question answering. *Cognitive Psychology*, 14(1), 78–106.
- Lewandowski, D., Kurowicka, D., & Joe, H. (2009). Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis*, 100(9), 1989–2001.
- Lewis, S., & Phillips, C. (2015). Aligning grammatical theories and language processing models. *Journal of Psycholinguistic Research*, 44, 27–46.
- Mahowald, K., James, A., Futrell, R., & Gibson, E. (2016). A meta-analysis of syntactic priming in language production. *Journal of Memory and Language*, 91, 5–27.
- Merchant, J. (2001). *The syntax of silence: Sluicing, islands, and the theory of ellipsis*. Oxford University Press.
- Momma, S. (2022). Producing filler-gap dependencies: structural priming evidence for two distinct combinatorial processes in production. *Journal of Memory and Language*.
- Momma, S., & Phillips, C. (2018). The relationship between parsing and generation. *Annual Review of Linguistics*, 4, 233–254.
- Müller, S., Abeillé, A., Borsley, R. D., & Koenig, J.-P. (2021). *Head-driven phrase structure grammar: The handbook (volume 9)*. Language Science Press.
- Newmeyer, F. J. (2017). Form and function in the evolution of grammar. *Cognitive Science*, 41, 259–276.
- Partee, B. H. (2014). A brief history of the syntax-semantics interface in western formal linguistics. *Semantics-Syntax Interface*, 1(1), 1–21.
- Pesetsky, D. (1995). *Zero syntax: Experiencers and cascades*. Cambridge, MA: MIT Press.
- Phillips, C., & Wagers, M. (2007). Relating structure and time in linguistics and psycholinguistics. In G. Gaskell & J. Altmann (Eds.), *Oxford handbook of psycholinguistics* (pp. 739–756). Oxford University Press.
- Pickering, M. J. (1993). Direct association and sentence processing: A reply to gorrell and to gibson and hickok. *Language and Cognitive Processes*, 8(2), 163–196.
- Pickering, M. J., & Barry, G. (1991). Sentence processing without empty categories. *Language and cognitive processes*, 6(3), 229–259.
- Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language*, 39(4), 633–651.
- Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: a critical review. *Psychological Bulletin*, 134(3), 427.
- Polinsky, M. (2013). Raising and control. In M. den Dikken (Ed.), *The cambridge handbook of generative syntax*. Cambridge, UK: Cambridge University Press.
- Pollard, C., & Sag, I. A. (1994). *Head-driven phrase structure grammar*. University of Chicago Press.
- Prasad, G., Van Schijndel, M., & Linzen, T. (2019). Using priming to uncover the organization of syntactic representations in neural language models. *arXiv preprint*

- arXiv:1909.10579*.
- R Core Team. (2021). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from [Mhttps://www.R-project.org/](https://www.R-project.org/)
- Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., Sutskever, I., et al. (2019). Language models are unsupervised multitask learners. *OpenAI blog*, 1(8), 9.
- Reitter, D., Keller, F., & Moore, J. D. (2011). A computational cognitive model of syntactic priming. *Cognitive Science*, 35(4), 587–637.
- Rizzi, L., & Cinque, G. (2016). Functional categories and syntactic theory. *Annual Review of Linguistics*, 2, 139–163.
- Roeper, T. (1987). Implicit arguments and the head-complement relation. *Linguistic inquiry*, 18(2), 267–310.
- Rosenbaum, P. S. (1965). *The grammar of english predicate complement constructions*. (Unpublished doctoral dissertation). Massachusetts Institute of Technology.
- Sag, I. A. (1976). *Deletion and logical form*. (Unpublished doctoral dissertation). Massachusetts Institute of Technology.
- Sag, I. A., & Fodor, J. D. (1994). Extraction without traces. In *West coast conference on formal linguistics* (Vol. 13, pp. 365–384).
- Schrimpf, M., Blank, I. A., Tuckute, G., Kauf, C., Hosseini, E. A., Kanwisher, N., ... Fedorenko, E. (2021). The neural architecture of language: Integrative modeling converges on predictive processing. *Proceedings of the National Academy of Sciences*, 118(45), e2105646118.
- Sinclair, A., Jumelet, J., Zuidema, W., & Fernández, R. (2022). Structural persistence in language models: Priming as a window into abstract language representations. *Transactions of the Association for Computational Linguistics*, 10, 1031–1050.
- Steedman, M. (2001). *The syntactic process*. Cambridge, MA: MIT press.
- Stowell, T. A. (1981). *Origins of phrase structure* (Unpublished doctoral dissertation). Massachusetts Institute of Technology.
- Williams, E. S. (1977). Discourse and logical form. *Linguistic inquiry*, 101–139.
- Wu, D. (2022). *Syntax and prosody of coordination* (Unpublished doctoral dissertation). Massachusetts Institute of Technology.
- Zehr, J., & Schwarz, F. (2018). *PennController for internet based experiments (IBEX)*. [Mhttps://doi.org/10.17605/OSF.IO/MD832](https://doi.org/10.17605/OSF.IO/MD832).
- Zhang, Y., Sun, S., Galley, M., Chen, Y.-C., Brockett, C., Gao, X., ... Dolan, B. (2019). Dialogpt: Large-scale generative pre-training for conversational response generation. *arXiv preprint arXiv:1911.00536*.
- Ziegler, J., Bencini, G., Goldberg, A., & Snedeker, J. (2019). How abstract is syntax? evidence from structural priming. *Cognition*, 193, 104045.