COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 43 (2019) e12749

ISSN: 1551-6709 online
DOI: 10.1111/cogs.12749

# Self-Priming in Production: Evidence for a Hybrid Model of Syntactic Priming

Cassandra L. Jacobs,[a] Sun-Joo Cho,[b] Duane G. Watson[b]

[a]*Department of Psychology, University of California*
[b]*Department of Psychology and Human Development, Vanderbilt University*

Received 17 January 2018; received in revised form 3 March 2019; accepted 9 May 2019

**Abstract**

Syntactic priming in language production is the increased likelihood of using a recently encountered syntactic structure. In this paper, we examine two theories of why speakers can be primed: error-driven learning accounts (Bock, Dell, Chang, & Onishi, 2007; Chang, Dell, & Bock, 2006) and activation-based accounts (Pickering & Branigan, 1999; Reitter, Keller, & Moore, 2011). Both theories predict that speakers should be primed by the syntactic choices of others, but only activation-based accounts predict that speakers should be able to prime themselves. Here we test whether speakers can be primed by their own productions in three behavioral experiments and find evidence of structural persistence following both comprehension and speakers' own productions. We also find that comprehension-based priming effects are larger for rarer syntactic structures than for more common ones, which is most consistent with error-driven accounts. Because neither error-driven accounts nor activation-based accounts fully explain the data, we propose a hybrid model.

*Keywords:* Syntactic priming; Structural persistence; Implicit learning; Language production

## 1. Introduction

Speakers typically reuse words and phrases that have already occurred in the conversation. There are a wide range of explanations for this type of repetition, which is also called priming. Researchers have claimed that priming serves as evidence that there is an abstract syntactic representation for production (Bock, 1986); that speakers are biased towards highly available representations (Bock & Loebell, 1990); that speakers and listeners attempt to align linguistic representations in order to facilitate communication (Pickering & Garrod, 2004, 2013); that speakers are implicitly learning statistical regularities in

Correspondence should be sent to Cassandra L. Jacobs, Department of Psychology, University of California, 135 Young Hall, One Shields Avenue, Davis, CA 95616. E-mail: clxjacobs@ucdavis.edu

their language and are attempting to match those statistics (Chang et al., 2006; Fine & Jaeger, 2013; Jaeger & Snider, 2013); and that speakers use the same linguistic structures as their listeners to signal social compatibility (Babcock, Ta, & Ickes, 2014; Ireland et al., 2011). Some of the above theories argue that producing a structure and hearing a structure should result in similar levels of priming because comprehension and production share syntactic representations (Pickering & Garrod, 2013; Tooley & Bock, 2014).

However, although there is a great deal of work demonstrating that speakers are primed by the syntactic structures that they comprehend, there is less evidence that speakers are primed by their own syntactic productions. Several extant theories and computational models of language processing predict that speakers should prime themselves (Jaeger & Snider, 2013; Pickering & Garrod, 2004, 2013; Reitter et al., 2011), most notably activation-based accounts of language production. By contrast, theories that state that priming requires an error signal (e.g., Chang et al., 2006) predict that speakers should not prime themselves because their produced utterance does not typically differ from the intended one (except in the case of speech errors). Despite the clear theoretical value of identifying whether self-priming occurs, there is relatively little experimental work looking at self-priming directly. The goal of the present work is to investigate whether speakers can prime themselves.

We explore this question in the domain of syntactic production (Bernolet & Hartsuiker, 2010; Bock, 1986; Bock & Griffin, 2000; Hartsuiker, Bernolet, Schoonbaert, Speybroeck, & Vanderelst, 2008; Jaeger & Snider, 2013; Kaschak, 2007; Reitter et al., 2011; Tooley & Bock, 2014). In classic syntactic priming studies (Bock, 1986; Bock & Griffin, 2000), participants hear and repeat descriptions of scenes in which the pictures could be described using two syntactic structures, such as the passive or active form of a transitive verb, or the double object or prepositional dative form of a ditransitive verb. For example, in Bock (1986) participants heard either an active or passive description of a scene, such as “Lightning struck the church” or “The church was struck by lightning." Bock (1986) found that when describing a new scene, participants were more likely to re-use the structure they had just heard than to use the alternate structure.

We refer to the type of priming that results from hearing or repeating a syntactic structure as comprehension-to-production priming. This type of priming occurs in a variety of contexts, including tasks in which speakers are asked to remember what they just heard (Bock, 1986), tasks where listeners hear descriptions of scenes generated by other people (Bock et al., 2007), where they complete sentence fragments (Kaschak, 2007; Kaschak, Kutta, & Coyle, 2014), and where they are talking with other people (Reitter & Moore, 2014). This type of priming has been shown to be robust in the face of a number of different factors (Mahowald, James, Futrell, & Gibson, 2016).

There are two prevailing classes of theories that explain syntactic priming. The first are error-based models (Chang et al., 2006; Dell & Chang, 2014; Fine & Jaeger, 2013; Jaeger & Snider, 2013). Although the various error-based models differ in their implementation, at a computational level they propose that syntactic representations in adults and children are constructed such that they match the statistics of syntactic structures in their environment, and such that more unexpected linguistic outcomes lead to greater learning. When a listener or reader encounters a syntactic structure that is unexpected, an error signal is generated that leads the listener to adjust their expectations for what syntactic structures will occur in the future. The less expected a construction is, the larger the error signal will be, and the greater the change in speakers' expectations. For example, a listener who encounters a less frequent construction, like a passive sentence (e.g., "The church was struck by lightning") will be more primed than they would be if they had encountered a more frequent construction, like an active sentence (e.g., "The lightning struck the church"). This is because the less frequent passive structure generates a larger error signal when expectations are violated, resulting in greater priming. This particular phenomenon is known as the inverse frequency effect and is well-documented in the syntactic priming literature across a wide variety of structures (Bock, 1986; Chang et al., 2006; Ferreira, 2003; Hartsuiker & Westenberg, 2000; Jaeger & Snider, 2013; Kaschak, Kutta, & Schatschneider, 2011; Mahowald et al., 2016).

The second class of theories explains comprehension-to-production priming by using spreading activation models (Bock, 1986; Hartsuiker et al., 2008; Pickering & Branigan, 1999; Reitter et al., 2011). Within these frameworks, the syntactic representations used by production and comprehension processes are shared, and priming is explained by increased activation of words and their associated syntactic structures in the language processing system. Thus, either comprehension or production experience with a particular syntactic structure in comprehension or production increases the activation or availability of a syntactic structure. After hearing a given construction, speakers are predicted to be more likely to produce that construction in the future, that is, demonstrating comprehension-to-production priming. Activation-based mechanisms do not typically account for the inverse frequency effect, though models like Reitter et al. (2011) have demonstrated that activation-based models demonstrate some sensitivity to syntactic structural probabilities.

Although both activation and error-based models can account for comprehension-to-production priming, they make differing predictions about whether uttering a completely self-generated syntactic structure makes one more likely to use that structure in the future (self-priming). Error-based models predict that speakers should not be able to prime themselves; if a speaker is producing what she intends to say, no error-signal is generated, and a speaker's present choice of a given syntactic construction should have little to no bearing on the type of construction that they will produce in the future. By contrast, activation-based models predict that speakers *should* be able to prime themselves; producing a given syntactic structure makes that syntactic structure more active, making that syntactic structure more likely to be produced in the future. Previous work, which we review below, has attempted to answer the question of whether speakers can prime their own syntactic choices. However, as we point out, even these studies are not pure examples of the type of self-priming that we have just described.

Two experimental paradigms have been used to examine the influence of speakers' syntactic productions directly: the "traffic light" paradigm (Segaert, Kempen, Petersson, & Hagoort, 2013) and the sentence completion paradigm (Kaschak, 2007; Kaschak et al., 2014). In the traffic light paradigm, the experimenter has complete control over the order in which speakers name referents. In this paradigm, referents in a display are color-coded as to which one is to be mentioned, and in what order. In the case of the active-passive alternation, when presented with an image of lightning striking a church, the speaker is tasked with naming either the church or the lighting first. While the experimenter has complete control over the order in which referents are named, one downside to this paradigm is that participants are forced into a particular conceptualization of a scene, which may be akin to a comprehension process, or which may generate an error signal for the speaker, as it differs from what they would have said spontaneously. This is particularly likely if the speaker's preferred order conflicted with the order imposed by the experimenter. At the very least, this procedure does not reflect the seemingly unconscious nature of syntactic choice. We also note that Segaert et al. (2013) operationalized priming as the attenuation of brain activation, but they do not report the behavioral responses for their participants. Although it is reasonable to think that an attenuated brain response might reflect a primed syntactic structure, we cannot know whether speakers were actually priming themselves without behavioral data.

Another paradigm that has been used to investigate self-priming is the sentence completion task (e.g., Kaschak, 2007; Kaschak et al., 2014). In this task, participants are given the stem of an utterance like "The rock star sold the undercover cop/cocaine ___" and are required to fill in the blank (e.g., "cocaine" or "to the cop"), which induces participants to produce a particular syntactic structure. Kaschak et al. (2014) found that participants were more likely to go on and produce this structure on later trials after completing the prime. Critically, the Kaschak studies all involve a comprehension component. That is, similar to the traffic light paradigm, subjects are forced into a certain conceptualization of an event and a sentence, which may engage the language comprehension system. Because participants necessarily comprehended the first parts of these sentences, any apparent self-priming observed in these studies could have resulted without participants producing anything at all.

Taken together, the data from the experiments by Kaschak et al. (2014) and Segaert et al. (2013) are suggestive, but they are not conclusive evidence of self-priming. In contrast to the experimental literature, corpus studies of spontaneous speech have often reported evidence for self-priming in syntactic production, but even this is mixed (Gries, 2005; Healey, Purver, & Howes, 2014; Jaeger & Snider, 2013; Reitter et al., 2011; Szmrecsanyi, 2008). Almost all of these studies have reported self-priming within a single speaker as well as priming from a conversational partner.

In a seminal corpus study, Gries (2005) found that speakers were much more likely to produce a structure that they had produced in a preceding utterance than reuse one that another speaker had produced while controlling for a number of potential confounds. In a separate study, Reitter and Moore (2014) characterized structural repetition as self-priming, and found that when ignoring repeated phrases, speakers were still highly likely to produce structures they had recently produced, and that the influence of one's own productions persisted for a longer period of time. Unfortunately, naturalistic corpus data cannot control for discourse constraints or speakers' preferences. For example, within a conversation, a certain conceptualization of discourse referents may lead to repeatedly using a syntactic choice that reflects that conceptualization, or speakers may have their own biases about which structure they prefer to use. Thus, while corpus studies are valuable, there are challenges in interpreting the repeated use of a syntactic structure across a conversation.

Surprisingly, to our knowledge, there are no experimental studies that have explicitly assessed self-priming in a task without a comprehension component, while still allowing speakers to make spontaneous choices about what to name. A task that combines the strengths of observational corpus linguistic studies with highly controlled materials designed to elicit particular target structures can help us address whether speakers' own *spontaneous* choices influence their later choices. Thus, an experimental investigation that removes potential confounds can serve as a useful complement to corpus-based and prior experimental work.

In addition to this, we hope to determine whether the syntactic choices of one production trial directly influence the syntactic choices of subsequent trials. While a number of studies have shown that speakers tend to repeat themselves, potentially due to self-priming or residual activation (e.g., Myslín & Levy, 2016; Reitter & Moore, 2014), these effects could result from confounds such as discourse factors favoring one structure over another, or a speaker's baseline preference for a construction, as we mention above. In order to determine whether there are dependencies between trials in the likelihood of selecting one structure over another, we apply a novel mixed effects autocorrelation analysis with a multinomial outcome to the current data set.

### *1.1. Self-priming as autocorrelation*

Autocorrelation is one way of conceptualizing dependencies between observations at two different points in time (Box & Jenkins, 1976; Cho, Brown-Schmidt, & Lee, 2018; Seedorff, Oleson, & McMurray, 2018). For example, the price of a single stock on the stock market today is highly predicted by the price it was yesterday. That is, the values associated with previous observations (say, time t1) may be correlated with later observations (e.g., t2, t3, . . ., tk) because their values are dependent on each other. This means that if we want to predict the price of a stock tomorrow, the value of the stock today is highly informative. Conversely, if we are interested in how some other factor influences the stock market tomorrow, we might want to control for today's stock market price. While typically autocorrelation is treated as a controlling factor to obtain unbiased estimators of interest (e.g., experimental condition effects), in the present study we are interested in assessing the degree of autocorrelation within a speaker as an indicator of syntactic self-priming. That is, if speakers' productions are dependent on recent productions and they prefer to reuse the same syntactic structures rather than switch to alternative structures, then there should be significant autocorrelations between time points within a speaker.

In three three-part experiments, we assessed comprehension-to-production priming as well as self-priming. The first phase of the experiment was a production task, in which we measured speakers' productions of the dative alternation (i.e. "the woman gave the book to the boy" vs. "the woman gave the boy the book") without any exposure to the construction from another speaker. If speakers are sensitive to their own productions (self-priming), then we should see dependencies between time points in syntactic choice above and beyond syntactic biases. Then, in the second phase of the experiment after the initial production task, participants listened to descriptions of ditransitive events containing only one syntactic form of the dative alternation in a comprehension-based manipulation. The third and final phase was again a production task, in which we measured the production rates of the primed structure in order to determine the effect of the comprehension priming phase. Thus, we tested whether speakers primed themselves and whether they showed comprehension-to-production priming. In Experiment 1, we used multiple prime types in the comprehension phase, and so we were also able to test whether the frequency of a given syntactic structure influenced the size of comprehension-to-production priming, providing a test of the inverse frequency effect.

All theories predict that exposure to a syntactic structure will bias speakers to produce more of that structure later. The two accounts make disjoint predictions about self-priming, however. Error-based accounts of priming predict that there should be no self-priming. These theories also predict an inverse frequency effect (Bock, 1986; Ferreira, 2003; Hartsuiker & Westenberg, 2000; Jaeger & Snider, 2013). The more infrequent a primed structure, the larger the effect of the prime should be on speakers (e.g., Chang et al., 2006; Jaeger & Snider, 2013). Because a lower probability structure leads to greater violations of expectation, this creates a larger error signal, leading to greater change in the production system. Activation-based models of priming (e.g., Pickering & Branigan, 1999; Reitter et al., 2011) predict that speakers *will* be affected by their own recent productions (i.e. self-priming), as well as by productions of others, but do not explain inverse frequency effects as readily.

We conducted three experiments to test whether speakers self-prime. The experiments contained either six primes (Experiment 1) or one (Experiment 2, 3a, 3b) comprehension prime. Experiment 1 manipulated the syntactic form of the prime (either a prepositional object, PO; a double object, DO; or an empty passive, the control). Experiments 2 and 3 use only the DO construction as the prime structure during the comprehension phase. Experiment 3a and 3b were conducted to further test for the self-priming effect, as well as replicate the single prime comprehension-to-production effect of Experiment 2. In all three experiments we test for comprehension-to-production priming and self-priming. We also test for inverse frequency effects in Experiment 1, as we could directly compare the effect of different prime types to each other.

## 2. Experiment 1

### 2.1. Methods

#### 2.1.1. Participants

Six hundred participants were recruited from Amazon Mechanical Turk. To ensure high reliability, workers were required to have a 99% HIT approval rating and to have completed at least 5,000 HITs. Participants were self-reported native speakers of English living in the United States. Each person received $1 USD for participation.

#### 2.1.2. Design

The experiment included three phases, an initial production phase, a comprehension phase in which participants were exposed to a prime structure, and a final production phase. There were three comprehension conditions. In the experimental comprehension phases, participants heard only a single syntactic construction. In the dative prime conditions, participants heard either double object (DO) constructions ("The waitress shows the customer the drinks") or prepositional object (PO) constructions for six ditransitive events ("The waitress shows the drinks to the customer"). In the control condition, participants were not exposed to the dative alternation; instead, they listened to three passive descriptions without prepositional phrases and three active (transitive) descriptions (e.g., "The man was stung"/ "The bee stung the man"). Stimuli for this experiment as well as Experiments 2 and 3 are presented in Appendix A.

#### 2.1.3. Materials

Twenty images depicting ditransitive events were taken from Bock and Loebell (1990). Fourteen of these were used for the two production phases (seven each), and six were used in the comprehension phase. An additional 30 filler images depicted non-ditransitive events. For the comprehension materials, a male speaker from Iowa provided descriptions of these six filler images. These recordings were spoken at a natural pace in a sound-attenuated booth. Recording order was randomized. We provide transcriptions of the recordings and prototypical descriptions of the images as part of Appendix A in Tables 1–4. We provide the recordings and images as supplemental material on the Open Science Framework.[1]

#### 2.1.4. Procedure

After the consent page, the experiment was presented as a three-page Qualtrics survey, with each page representing a separate phase in the experiment. Each page contained a fixed set of images whose order was randomized for each participant. The order of the production phases was counterbalanced across subjects. Participants could not return to any prior phase of the experiment and were required to give a response to all items to proceed but could quit at any time without consequence. Participants took 12.4 of the 25 min allotted on average to complete all three phases of the experiment.

The experiment was broken into three phases, which we depict in Fig. 1. Participants provided spontaneous written descriptions of the pictures on the first and third pages. In the first and third phases of the experiment, participants were told only to describe the picture and were not given any particular guidelines other than to write a simple description of the image. Because prior research has established that comprehension-to-production manipulations are effective in written production as well as spoken production (Branigan, Pickering, & Cleland, 1999; Hartsuiker et al., 2008; Pickering & Branigan, 1999), we used a typed sentence production task to test for syntactic persistence, which enabled us to conduct the experiment online. Participants typed descriptions of 7 pictures of ditransitive events, which included images like a man passing a pitcher to a woman or children showing a painting to their teacher. There were also 10 filler image descriptions. We provide example descriptions of the images in Table 1 of Appendix A.

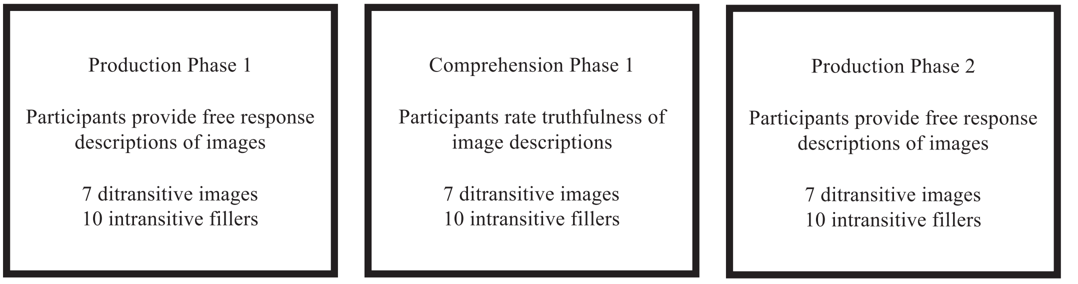


Fig. 1. Schematic of the timeline of three-phase experiment.

During the second (comprehension) phase of the experiment, participants were asked to evaluate descriptions of images. Participants rated how much the image corresponded to the recorded description on a scale of 0 to 100 with 100 being a perfect correspondence. In the control condition, participants heard only descriptions of transitive events in the active and passive forms without prepositional phrases (Table 3, Appendix A). In the experimental conditions, they heard either only double object or prepositional object descriptions of six ditransitive events (Table 2, Appendix A). Three filler pictures appeared with inappropriate descriptions and served as catch trials (Table 4, Appendix A). Participants could play the recordings as many times as needed to complete the task.

Participants then completed the second production task, which was identical to the first phase, but with a new set of images.

### 2.2. Results

We excluded participants if their ratings of the catch trials were above 60 out of 100 or if any ratings for the critical comprehension trials were under 40; their responses were ungrammatical (e.g., "She give a glass to boy"), but not telegraphic (e.g., "Girl gives glass to boy"); or more than two of seven critical responses was not a reasonable description of the image (e.g., "playing catch" to describe a dinner scene). We also excluded any participant who completed the experiment twice. Out of 600 participants, 498 remained after these exclusions. The DO, PO, and control conditions contained 163, 179, and 156 participants, respectively. The first author manually annotated which construction the participant used (double object [*DO*], prepositional object [*PO*], or *Other*) for all critical ditransitive event descriptions but was blind to prime form, trial number, and block number. On average, DO, PO, and Other utterances comprised 25.2%, 30.8%, and 44.0% of all descriptions on the first critical trial, respectively, and 26.3% ($SE$ = 0.02), 34.3% ($SE$ = 0.02), and 39.5% ($SE$ = 0.02) during the first phase overall.

#### 2.2.1. Effect of primes during the comprehension phase

First, we tested for the effect of the prime during the comprehension phase on participants' preferences toward producing the primed structure. To compare productions across all three conditions, we examined whether the likelihood of using a PO structure changed depending on the syntactic form of the prime (i.e. either DO, PO, or "control"). We selected the PO structure as it allows for us to test for the effects of PO priming and DO priming simultaneously. Relative to the control conditions, PO usage should drop following DO primes because the two structures compete for selection, as the DO structure will become more available for production (Bock, 1986). Similarly, when participants hear PO primes, PO usage should rise and DO usage should fall. We observed a rise in the use of the PO structure in the PO prime condition (31.3% to 40.3%), a decline in usage of the PO structure in the DO prime condition (from 36.9% to 24.7%; similarly, DOs rose from 24.1% to 42.5%), and a slight drop in PO usage in the passive prime condition (34.7% to 32.2%).

This analysis employed mixed effects models to test the probability that participants used a particular syntactic structure on a given trial. This model contained intercept and slope terms for participants but only a slope for the effect of comprehension input and an intercept for items, making it a maximal model. To test for the comprehension-to-production priming effect, the model included fixed effects for production phase (before/after comprehension input, difference coded) and the syntactic form of the prime (dummy coded with the passive control as the baseline). These allowed us to include any potential baseline differences between the conditions. The critical analysis for comprehension impacting production is the interaction between these two terms.

We found that relative to the control condition, PO usage falls in the DO prime condition (because DO usage rises), and PO usage rises in the PO prime condition. These data are consistent with comprehension-to-production priming effects. We can only speculate that the marginal decrease in PO usage in the control condition may be due to the intentional omission of "by" passives (Bock & Loebell, 1990). The fixed effect results are presented in Table 1 below. In Fig. 2, we plot the comprehension priming effect in addition to trial-by-trial PO usage. Priming can be seen in the discontinuity from production trial 7 (production phase 1) to 8 (production phase 2).

#### 2.2.2. *Influence of prior structural probabilities on the influence of comprehended primes*

Error-driven learning accounts predict that speakers will show more priming in response to a rarer syntactic structure compared to a more frequent one (e.g., Bock, 1986;

Table 1
Fixed effects results for the logit mixed model analysis of priming (changes in prepositional object (PO) production) in all conditions from Experiment 1

| Fixed Effect | Coding | Estimate | *SE* | Wald *z* | *p* |
|---|---|---|---|---|---|
| Intercept | | −0.99 | 0.21 | −4.66 | * |
| Before/After | Difference (−1/1) | −0.13 | 0.07 | −1.78 | . |
| Double object (DO) baseline | Dummy (passive baseline) | −0.19 | 0.16 | −1.19 | |
| PO baseline | Dummy (passive baseline) | 0.16 | 0.16 | 1.02 | |
| DO × Before/After | Dummy × Difference | −0.33 | 0.09 | −3.66 | *** |
| PO × Before/After | Dummy × Difference | 0.36 | 0.08 | 4.19 | *** |

*Note*. "." Indicates marginal significance at, *indicates significance at <.05, ***at <.001.

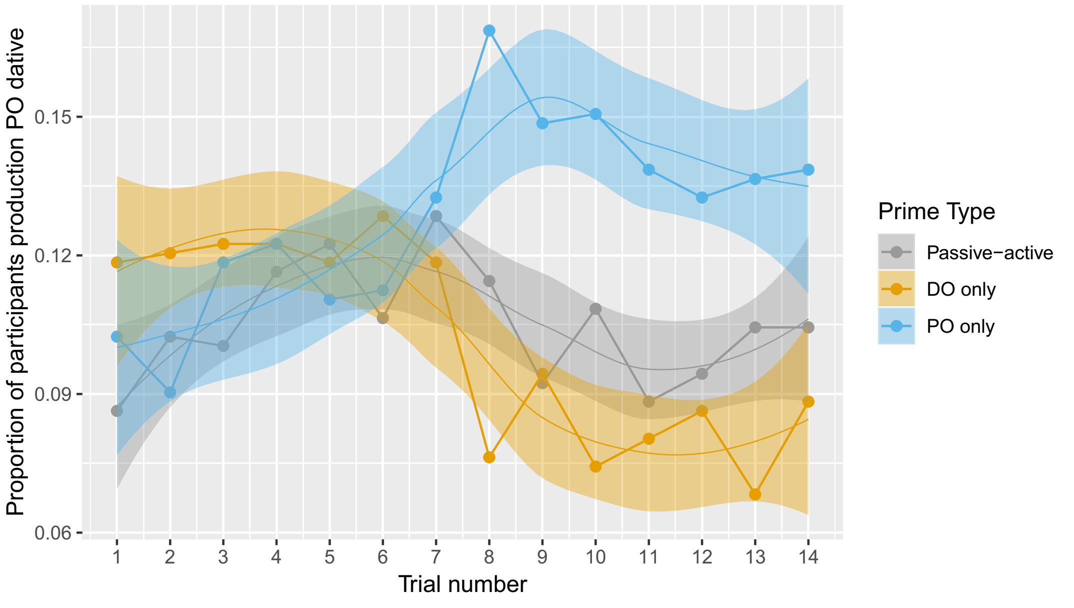


Fig. 2. Proportion of trials containing the prepositional object (PO) structure in Experiment 1 by condition. Error regions represent locally computed bootstrapped moving window averages.

Chang et al., 2006; Ferreira, 2003; Hartsuiker & Westenberg, 2000; Jaeger & Snider, 2013; Kaschak et al., 2011; Mahowald et al., 2016). To assess whether speakers were more strongly primed by one syntactic structure or another, we recoded the variable representing the syntactic structure the speaker selected on each trial as either the same as the prime form (1) they would later receive or different (−1). We excluded the control condition from the analyses here. For the DO prime condition, we were assessing the magnitude of increase in the use of DOs across the first and second production phases; for the PO prime condition, we tested for the likelihood that speakers would use more POs in the second production phase than in the first. We specifically predicted that, if there is an inverse frequency effect, then we expect greater priming for the lower-probability structure (DO) than for the higher-probability one (PO). All variables were centered and scaled.

To test for differential priming effects, we constructed a logit mixed effects model that allowed us to compare the DO and PO priming conditions directly. Specifically, we modeled whether the structure that would eventually be primed was more likely to be used in the second half (Before/After) depending on the prime type (DOs/POs), and their interaction, the effect of interest. A significant interaction between prime type and production phase would suggest that DO and PO primes impact speakers to different extents. We visualize the increase in DO usage in the DO prime condition and PO usage in the PO prime condition together in Fig. 3 below.

The statistical model demonstrated that speakers were indeed primed significantly more by the DO prime than the PO prime, even though both led to greater use of the primed

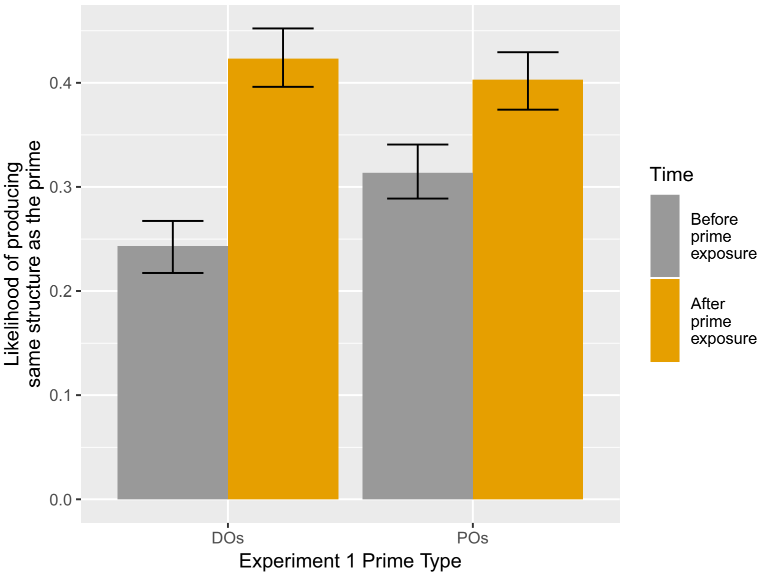


Fig. 3. Effect of prime exposure during the comprehension phase on structure use in the second production phase. Both double object (DO) and prepositional object (PO) primes lead to more DO and PO use, respectively, but the increase in use of DO structures following DO primes is larger than the increase for the PO structure. Error bars represent bootstrapped standard errors of the mean.

structure in the second phase of the experiment (after prime exposure) than before. We provide the fixed effects estimates of this analysis below in Table 2.

The results here demonstrate that DO priming led to even greater changes in speakers' syntactic preferences than PO priming, consistent with error-driven learning accounts of syntactic priming (e.g., Chang et al., 2006), and consistent with many prior studies demonstrating the inverse frequency effect (Bock, 1986; Chang et al., 2006; Ferreira, 2003; Hartsuiker & Westenberg, 2000; Jaeger & Snider, 2013; Kaschak et al., 2011; Mahowald et al., 2016).

### 2.2.3. *Self-priming*

Earlier, we operationalized self-priming as the tendency for speakers to reuse a syntactic structure beyond what one would expect given the base rate of the syntactic structure alone. For visualization of rates of structural reuse, we partitioned the first phase's production data into pairs of productions. These pairs represent the possible types of trial-to-trial dependencies. For example, on trial 1, a speaker might produce a PO, DO, or "Other." Similarly, on trial 2, they can make the same set of decisions. If self-priming can be thought of as the tendency to repeat or reuse a syntactic structure, then we should see greater chances of DO-DO, PO-PO, and Other-Other outcomes than "switch" outcomes (DO-PO, DO-Other, PO-DO, PO-Other, Other-DO, Other-PO). For visualization we computed the these *transition probabilities* as the likelihood that DOs, POs, and Others would follow the three different utterance types (DO, PO, Other). To calculate

Table 2
Fixed effects results for the logit mixed model analysis combining double object (DO) and prepositional object (PO) conditions from Experiment 1

| Fixed Effect | Coding | Estimate | *SE* | Wald *z* | *p* |
|---|---|---|---|---|---|
| Intercept | | −0.87 | 0.16 | −5.39 | *** |
| Before/After | Difference (−1/1) | 0.81 | 0.11 | 7.01 | *** |
| Experiment (PO baseline) | Difference (−1/1) | −0.07 | 0.07 | −1.08 | |
| Experiment × Before/After | Difference × Difference | 0.28 | 0.08 | 3.39 | *** |

*Note*. ***Indicates significance at <.001.

these transition probabilities, we counted the proportions of the different outcomes across all pairs of trials (i.e., 1–2, 2–3, …, 6–7) in the first production phase. Visual inspection shows that speakers were highly likely to reuse the most recently produced syntactic structure. We present the tendency for speakers to use different pairs of syntactic structures across trials in the first production phase below in Fig. 4. Repetition was the most likely outcome of all three types of transitions: given that speakers produced a DO, Other, or PO on the previous trial, speakers tended to repeat DO structures 40.6% of the time, Other structures 47% of the time, and PO structures 46% of the time, respectively.

To determine whether there were dependencies between trials, we analyzed the data using an autocorrelation to determine the dependencies between speakers' individual productions in a time series, as described in the Introduction. While this technique is rarely applied to psycholinguistic data (with a few exceptions: Cho et al., 2018; Seedorff et al., 2018), it is ideal for determining whether speakers engage in syntactic self-priming. Here we apply this analysis to the first production phase only, as that allows for a pure determination of self-priming without contamination from the comprehension phase. A post-hoc analysis of the second production phase showed the same pattern of results—which we take as evidence that speakers prime themselves—as in the first production phase. These results are presented in Table 2 of Appendix B.

Speakers' syntactic decisions are discrete, unordered categorical outcomes, so the model we construct is therefore a mixed-effects *multinomial* logistic model (Hedeker, 2003; Skrondal & Rabe-Hesketh, 2003). Using a mixed-effects multinomial logistic approach is important because speakers can make any of three decisions on each critical trial (DO, PO, or Other). We extended existing mixed-effect multinomial logistic approaches to model transitions between different discrete outcomes, as others have done before in applications such as employment status (e.g., Pettitt, Tran, Haynes, & Hay, 2006). The extended model, called a Markov mixed-effect multinomial logistic model, assesses self-priming by estimating the relationship between categorical outcomes on pairs of trials at two adjacent points in time (i.e., the previous trial and the current one), which we illustrate in Fig. 5. Modeling multinomial data is not currently feasible in the lme4 package in R (Bates, Maechler, Bolker, & Walker, 2014), so we used WinBUGS 1.4.3 (Spiegelhalter, Thomas, & Best, 2003), a Bayesian analysis package that uses Markov chain Monte Carlo (MCMC) methods for parameter estimation. We provide model specifics in Appendix C.

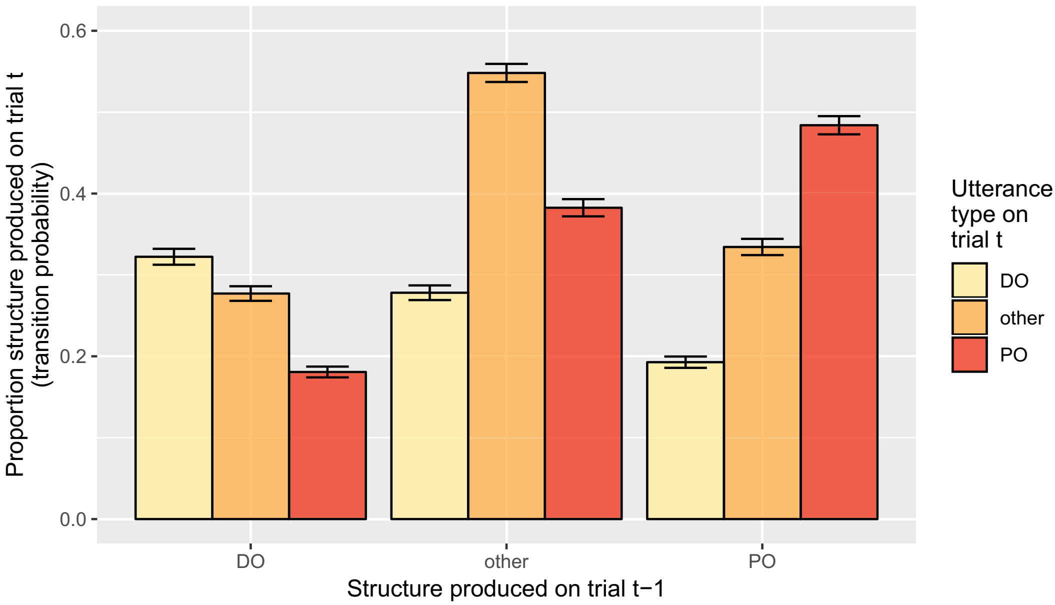


Fig. 4. Transition probabilities between syntactic structures across trials, Experiment 1, first production phase. Speakers repeat the structure they have just produced more often than switching to produce either different structure.

| | | Current trial | | |
|---|---|---|---|---|
| | | Other | DO | PO |
| Previous trial | Other | | | |
| | DO | | | |
| | PO | | | |

Fig. 5. Possible pairs of syntactic structures chosen on two adjacent time points. Gray squares represent syntactic structure re-use; that is, a first-order autocorrelation between time steps (trials).

Because we are ultimately interested in whether a speaker produces a DO, PO, or Other relative to the previous trial, we must incorporate whether the prior trial was DO, PO, or Other as a dependent variable, which we represent with dummy coding requiring two variables, with Other as baseline (0, 0), DO as (1, 0), and PO as (0, 1). Using the categorical outcome of only the previous observation as a dependent variable for the next trial is known as a *first-order autocorrelation (AR(1))*. Because there are two (dummy-coded) dependent variables (DO as (1, 0) and PO as (0, 1)) and two (dummy-coded) covariates (DO as (1, 0) and PO as (0, 1)), the model ultimately estimates four coefficients for the fixed AR(1) effects and four random effects to account for individual differences in AR(1) effects. These sets of parameters for both the fixed and random effects thus represent the four potential outcome types of interest over pairs of adjacent trials. Repetition is represented by DO-DO and PO-PO trial pairs; on the other hand, a *switch* in syntactic structures is evidenced by DO-PO and PO-DO pairs. Estimates in the model are on the logit scale.

In a Markov mixed-effect multinomial logistic model, we estimate six random effects, four random slopes and two random intercepts. For the random effects, the model estimates for each speaker whether that speaker uses the same structure (e.g., DO) on the current trial as on the previous trial, or instead switches to producing the alternate structure (e.g., PO) relative to their probability of producing an unrelated syntactic structure (e.g., Other). The two random intercept terms represent estimates of a speaker's bias toward Other. The four random slopes include two estimates of a speaker's likelihood of repeating the same syntactic structure (DO-DO, PO-PO outcomes); the other two estimates represent the likelihood that speakers will switch to the alternate structure over Other (DO-PO, PO-DO).

We use Other as a baseline and compare the probability of producing a DO versus Other and the probability of producing a PO versus Other. Evidence of syntactic self-priming arises when the estimated repetition (DO-DO, PO-PO) coefficients are positive, which indicates that a speaker reuses a particular structure more than they produce DO-Other and PO-Other pairs (controlling for the other effects in the model). When evaluating the model for autocorrelations that would support evidence of syntactic self-priming, it is also important that we find that the DO-PO and PO-DO (switch) estimates are *not significantly positive*, because significantly positive coefficients show that speakers choose one form of a syntactic structure over another, rather than simply producing more dative sentences. If both repetition and switch coefficients are positive, this suggests only that speakers are producing fewer Other utterances. As we used Bayesian models to estimate coefficients, we present credible intervals (CI) rather than confidence intervals. Further discussion of credible intervals can be found in Morey, Hoekstra, Rouder, Lee, and Wagenmakers (2016). CIs that include 0 imply that there is a not a significant effect.

Results of the Markov mixed-effect multinomial logistic model are presented in Table 3. We highlight the sections of the table in gray to illustrate the places where we expect positive (non-0) coefficients demonstrating the tendency for speakers to produce DO-DO pairs or PO-PO pairs across trials. In line with our predictions for self-priming stated above, both the DO-DO and PO-PO estimates are significantly positive. Speakers who produced a DO on the previous trial are 1.58 times more likely to produce a DO on the current trial. Similarly, speakers who produced a PO on the previous trial are 1.51 times more likely to produce a PO on the current trial. Importantly, DO-PO and PO-DO coefficients are not significant, showing that speakers are not simply learning to produce more datives in general. That is, we see a significant self-priming effect. We note that participants vary in the extent to which they exhibit self-priming, as shown by the variance estimates of the random effects, which we present in Table 1 in Appendix B.

### *2.3. Discussion*

Overall, we replicate previous priming findings. Speakers are more likely to produce a syntactic structure that they have recently heard. In addition, we find that the frequency

Table 3
Fixed effects of the Markov mixed-effect multinomial logistic regression model from Experiment 1

| | Fixed Effects, Trial 1 | | | | | |
|---|---|---|---|---|---|---|
| | Other Versus Double Object (DO) Outcome | | | Other Versus Prepositional Object (PO) Outcome | | |
| | Estimate | l-CI | u-CI | Estimate | l-CI | u-CI |
| Grand Mean | −0.44 | −2.32 | 0.55 | −0.58 | −2.81 | 0.68 |
| AR1 for other versus DO | | – | | | – | |
| AR1 for other versus PO | | – | | | – | |

| | Fixed Effects, Trials 2–7 | | | | | |
|---|---|---|---|---|---|---|
| | Other Versus Double Object (DO) Outcome | | | Other Versus Prepositional Object (PO) Outcome | | |
| | Estimate | l-CI | u-CI | Estimate | l-CI | u-CI |
| Grand Mean | −0.65 | −1.21 | 0.17 | −0.23 | −0.73 | 0.38 |
| AR1 for other versus DO | 0.47 | 0.05 | 0.90* | −0.26 | −0.65 | 0.10 |
| AR1 for other versus PO | −0.02 | −0.43 | 0.35 | 0.40 | 0.07 | 0.76* |

*Note*. CI indicates 95% Bayesian credible interval; "l-CI" indicates the lower limit and "u-CI" indicates the upper limit. Significance marked by * for fixed effects based on 95% Bayesian credible interval testing. "–" indicates the parameters are not estimated (because they do not exist at Trial 1).

of the structure is inversely correlated with the size of the priming effect. Finally, and most important, we see evidence that speakers prime themselves: choosing to produce a DO or PO makes producing that specific structure more likely in the future.

One issue that may puzzle the reader is the apparent discrepancy between evidence suggesting that speakers prime themselves and the absence of clear visual evidence for this in Fig. 2, which show the proportion of DO and PO trials as a function of trial number. There is no visual change in the proportion of either syntactic type across trials in the initial production phase of the experiment. This lack of a clear visual trend may be due to aggregating the data across the population. Individual speakers may be more likely to choose a given structure over trials, but this is lost when we average across all the speakers in this experiment. For example, imagine on trial 1 that 40% of speakers chose DO, 40% of speakers chose PO, and 10% chose Other. On trial 2, the DO biased speakers might be more likely to produce a DO and the PO speakers more likely to produce a PO, but the average across the population would not change. This highlights the importance of looking beyond aggregate descriptive statistics when studying self-priming.

One potential concern is that effects we see here are not due to syntactic priming, but instead are the result of conceptual priming; that is, apparent preferences to repeat the same structure are the result of priming the semantics of ditransitive verbs. We can rule out this possibility because the autocorrelation analysis allows us to determine whether speakers are simply learning that there are dative images in the experiment or whether speakers are truly priming themselves. As we see in Table 2, only the DO-DO and PO-

PO parameters are significant, demonstrating structure-specific self-priming, with no evidence that speakers are simply learning to avoid producing Other utterances. Rather, speakers are reusing the same syntactic structures, suggesting that the self-priming we see here is due to structure reuse and not just conceptual priming.

## 3. Experiment 2

As in Experiment 1, we were interested in examining the change in participants' syntactic preferences between the first production phase and the second production phase. The critical difference between Experiment 1 and 2 is that Experiment 2 only had a single comprehension prime. Experiment 2 was originally designed to create a context in which comprehension-to-production priming and self-priming could be directly compared (i.e. to examine the effects of a single syntactic choice on language production). However, this experiment was designed before the decision was made to conduct an auto-correlation analysis, which allows for a direct examination of trial-to-trial dependencies. Here we include Experiment 2 as a replication of the apparent self-priming effects in Experiment 1. Because Experiment 1 demonstrated that both DO structures and PO structures lead to priming, to simplify both the design and analysis in this experiment, we only tested comprehension-to-production priming for a single structure, DO, which we selected because priming with this structure yielded the largest priming effect in Experiment 1.

### *3.1. Methods*

#### *3.1.1. Participants*

Three hundred and sixty individuals took part in the study via Mechanical Turk using the same criteria as in Experiment 1. In addition, we excluded anyone who had previously participated in any part of Experiment 1 or who attempted to participate in this experiment multiple times, leaving 333 participants. Participants received $1 USD for participation in this experiment, which took approximately 12 min on average to complete out of the allotted 25 min.

#### *3.1.2. Counterbalance*

We took the six prime images from Experiment 1 and created six lists, each one containing only a single prime. We counterbalanced the order of the images occurring in the first and second production phases as we did for Experiment 1, creating a total of 12 lists.

#### *3.1.3. Procedure*

Other than the number of primes (1 vs. 6) in the comprehension phase of the experiment, the production phases of Experiment 2 were identical to Experiment 1. In the first production phase, participants described 7 ditransitive images plus 10 fillers. In the comprehension phase, participants rated the truthfulness of a single ditransitive description using the DO form along with 10 fillers, 3 of which were incorrect. In a second production phase, participants described an additional 7 ditransitive images along with 10 fillers. Participants saw all trials within a phase in a random order.

### 3.2. Results

As in the first experiment, we built maximal logistic mixed effects models assessing the influence of the comprehension prime on participants' tendencies to produce the primed structure with participants and items as random slopes and intercepts. In this analysis, we model the likelihood of a participant producing a DO structure. This is a change from Experiment 1, in which we analyzed the likelihood of a PO structure. Following Experiment 1, we predicted that participants would use more DOs after hearing a DO structure (a main effect of block). Self-priming would be evident in trial-to-trial changes in the likelihood of one structure or speakers' preferred structures.

#### 3.2.1. Comprehension-to-production priming

We conducted a single analysis to test for the influence of the comprehension prime on participants' production preferences, specifically predicting whether speakers were more likely to use DOs after hearing a DO prime. Replicating the results of Experiment 1, participants used more DOs in the second production phase (33.3%; *SE* = 0.03) than the first phase (25.6%; *SE* = 0.02), demonstrating robust priming effects in response to even a single comprehended DO prime.[2] These results are summarized in Table 4 below and plotted in Fig. 6.

#### 3.2.2. Self-priming

We again tested for the possibility that speakers might prime themselves. The Markov mixed-effect multinomial logistic regression model described in Experiment 1 (also see Appendix C) was fit to the data for the first seven trials. Table 5 shows the results of the model. The odds of producing a DO as compared to Other after having just producing a DO are 1.92 times higher than the odds of producing a DO after having produced an Other. Similarly, the odds of producing a PO compared to Other after having just produced a PO are 2.03 times higher than the odds of producing a PO after having produced an Other. Again, there are non-ignorable individual differences in self-priming effects. We present the estimates for the variances of random effects in Table 3 in Appendix B. A post-hoc analysis failed to find evidence that speakers prime themselves during the

Table 4
Fixed effects and results for the logit mixed model analysis from Experiment 2. Participants used significantly more DOs in the second half than in the first

| Fixed Effect | Coding | Estimate | *SE* | Wald *z* | *p* |
|---|---|---|---|---|---|
| Intercept | | −1.20 | 0.21 | −5.62 | *** |
| DO priming before/after | Difference | 0.26 | 0.06 | 4.38 | *** |

*Note*. *** indicates significance at < .001.

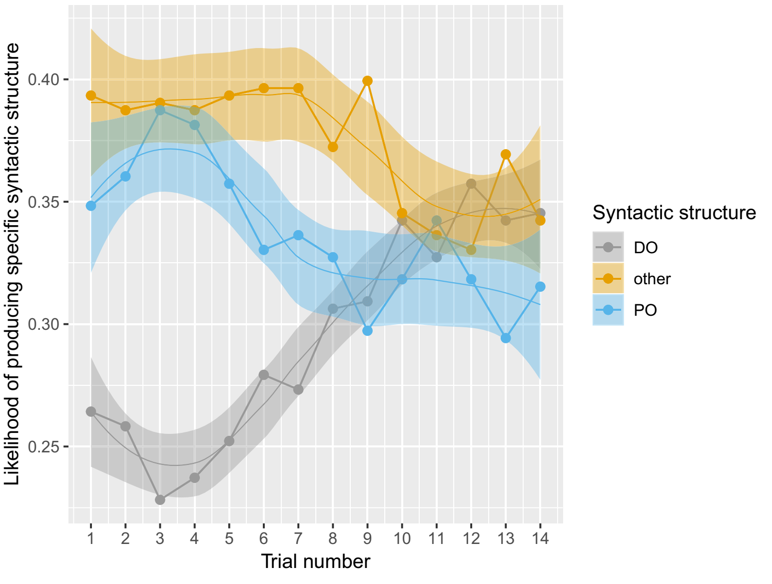


Fig. 6. Effect of comprehension of a single double object (DO) prime on DO, prepositional object (PO), and other production. Error regions represent locally computed bootstrapped moving window averages.

second production phase, perhaps due to contamination or ceiling effects from comprehension input. These results are presented in Table 4 of Appendix B.

### *3.3. Discussion*

Again, there are significantly variable effects of self-priming among participants in both production phases. CIs in both experiments are quite wide, suggesting high variability. The high variability of the effects here and in Experiment 1 merit further investigation.

One potential explanation of the apparent self-priming effects of Experiments 1 and 2 is that participants were always able to revise their responses before moving on to the next phase. While we do not know that speakers chose to edit their responses to make them stylistically consistent, we sought to rule out this possibility in Experiment 3 by changing the responses to occur on a single page, analogous to an experiment conducted in the laboratory. Finally, an additional goal of Experiment 3 was to replicate the comprehension-to-production priming effect resulting from exposure to a single prime, which was smaller in magnitude (7.7% vs. 18.4%) than Experiment 1.

## 4. Experiment 3

Due to the size of the self-priming effect in Experiments 1 and 2, as well as the surprising effectiveness of a single comprehension prime on production preferences, we conducted Experiment 3 to replicate the self-priming and comprehension-to-production

Table 5
Fixed effects of the Markov mixed-effect multinomial logistic regression model from Experiment 2

| | Fixed Effects, Trial 1 | | | | | |
|---|---|---|---|---|---|---|
| | Other Versus Double Object (DO) Outcome | | | Other Versus Prepositional Object (PO) Outcome | | |
| | Estimate | l-CI | u-CI | Estimate | l-CI | u-CI |
| Grand Mean | −0.70 | −1.65 | 0.18 | −0.18 | −0.97 | 0.50 |
| AR1 for other versus DO | | – | | | – | |
| AR1 for other versus PO | | – | | | – | |
| | **Fixed Effects, Trials 2–7** | | | | | |
| | Other Versus Double Object (DO) Outcome | | | Other Versus Prepositional Object (PO) Outcome | | |
| | Estimate | l-CI | u-CI | Estimate | l-CI | u-CI |
| Mean for other | −0.91 | −1.54 | −0.16* | −0.45 | −1.15 | 0.20 |
| AR1 for other versus DO | 0.64 | 0.10 | 1.06* | 0.25 | −0.18 | 0.66 |
| AR1 for other versus PO | 0.12 | −0.35 | 0.52 | 0.71 | 0.25 | 1.15* |

*Note*. CI indicates 95% Bayesian credible interval; "l-CI" indicates the lower limit and "u-CI" indicates the upper limit. Significance marked by * for fixed effects based on 95% Bayesian credible interval testing. "–" indicates the parameters are not estimated (because they do not exist at Trial 1).

effects of Experiment 2. We made an additional change to the experimental procedure; each response was presented one at a time, with no opportunity to revise responses by returning to a previous question, making the experiment more analogous to a laboratory experiment.

Experiment 3 is divided into two parts. In Experiment 3a, the critical dative items were not randomized in the production phase due to a design error. In Experiment 3b, the critical dative items were appropriately randomized. To foreshadow the results, we do not find evidence of self-priming in either experiment, but when their results are collapsed, we do find evidence for self-priming in both the first and second production phase. Below, we first present the findings from each study individually, and then the results of the studies collapsed in a post-hoc analysis.

### *4.1. Experiment 3a*

#### *4.1.1. Methods*

*4.1.1.1. Participants*: We recruited 360 participants via Amazon Mechanical Turk with the same constraints as Experiments 1 and 2. Participants were paid $2 USD for their participation in keeping with more recent pay standards for crowd workers. We excluded 40 participants whose responses were significantly truncated, clearly non-native or unnatural, or demonstrated inattention. This left us with 320 participants in total with acceptable responses.

*4.1.1.2. Materials*: The materials and experimental design of Experiment 3a were the same as in Experiment 2.

*4.1.1.3. Procedure*: Unlike Experiments 1 and 2, each question was presented on its own page. Participants saw the consent page, followed by instructions for each image to describe the image (with no further instruction). In the first production phase, participants produced 7 utterances describing ditransitive events, followed by 10 filler utterances, each on its own page. Then, participants rated 1 description containing the double object form of the dative alternation, followed by 10 filler descriptions, each on its own page. In the second production phase, participants again provided descriptions of 7 ditransitive event images, followed by 10 filler descriptions.

*4.1.2. Results*

*4.1.2.1. Coding*: The first author manually coded a subset of the experimental data. Then, in order to speed up the annotation process, a random forest classifier was trained on the text descriptions and syntactic structure outcomes of the manually coded subset of the data. The classifier then analyzed the remaining set of the data, and the coder manually checked the output of the entire set of automatically classified utterances. Of the labels, 11.6% needed to be corrected from the model-predicted outcome. As before, transcription was blind to production phase (before or after the comprehension phase) and counterbalance/prime information.

*4.1.2.2. Comprehension-to-production priming*: In the first production phase, prior to any external exposure to the dative construction of interest, participants produced the dative construction 63.2% of the time, with 25.7% of productions being the DO form ($SE$ = 0.02), and 37.5% ($SE$ = 0.03) being the PO form. In the second half of the experiment, participants produced the DO form 31.9% of the time ($SE$ = 0.03), and the PO form 32.2% of the time ($SE$ = 0.03), representing a drop from an 11.8% advantage to a 0.3% advantage for the PO form over the DO form of the alternation, and a 6.2% rise in DO usage.

Statistical analyses conducted with maximal mixed effects models (intercepts and slopes for the fixed effects for both participants and items) predicting whether speakers would produce the DO form of the dative alternation confirmed that speakers were more likely to produce a DO in the second production phase of the experiment. We present these statistical analyses below in Table 6.

*4.1.2.3. Self-priming (AR1 analysis)*: We tested whether speakers prime themselves using the same multivariate autocorrelation regression technique as in Experiments 1 and 2. Unlike Experiments 1 and 2, however, the first-order autocorrelation analysis of this experiment found no statistically significant effect demonstrating self-priming, though there was a slight tendency for speakers to produce fewer POs after DOs. We present the statistical estimates in Table 7 below and the random effects in Table 5 of Appendix B. A post-hoc analysis of the second production phase data also failed to find consistent

Table 6
Mixed effects model results of the effect of the comprehension prime on participants' likelihood of producing a Double Object (DO) structure

| Fixed Effect | Variable Type | Estimate | *SE* | z | *p* |
|---|---|---|---|---|---|
| Intercept | | −1.25 | 0.22 | −5.68 | *** |
| Before/After | Contrast coding | 0.13 | 0.06 | 2.17 | * |

*Note*. *indicates significance at $p < .05$. *** indicates significance at < .001.

Table 7
Fixed effects of the Markov mixed-effect multinomial logistic regression model from Experiment 3a

| | Fixed Effects, Trial 1 | | | | | |
|---|---|---|---|---|---|---|
| | Other Versus Double Object (DO) Outcome | | | Other Versus Prepositional Object (PO) Outcome | | |
| | Estimate | l-CI | u-CI | Estimate | l-CI | u-CI |
| Grand Mean | −0.13 | −1.95 | 3.39 | 1.69 | −1.29 | 5.03 |
| AR1 for other versus DO | | – | | | – | |
| AR1 for other versus PO | | – | | | – | |

| | Fixed Effects, Trials 2–7 | | | | | |
|---|---|---|---|---|---|---|
| | Other Versus Double Object (DO) Outcome | | | Other Versus Prepositional Object (PO) Outcome | | |
| | Estimate | l-CI | u-CI | Estimate | l-CI | u-CI |
| Mean for other | −0.23 | −0.98 | 0.56 | 0.04 | −0.51 | 0.76 |
| AR1 for other versus DO | −0.25 | −0.81 | 0.30 | −0.09 | −0.61 | 0.36 |
| AR1 for other versus PO | −0.55 | −1.06 | −0.07* | 0.22 | −0.31 | 0.64 |

*Note*. CI indicates 95% Bayesian credible interval; "l-CI" indicates the lower limit and "u-CI" indicates the upper limit. Significance marked by * for fixed effects based on 95% Bayesian credible interval testing. "–" indicates the parameters are not estimated (because they do not exist at Trial 1).

effects of self-priming, potentially due to contamination from comprehension input. We present the results of this analysis in Table 6 of Appendix B.

### 4.2. *Experiment 3b*

Experiment 3a replicated the comprehension-to-production priming effect with a single DO prime from Experiment 2 using a single page response format, rather than allowing participants to view all productions simultaneously. However, Experiment 3a failed to find a significant self-priming effect and trials were not randomized due to experimenter error. To correct for this, as well as to replicate the single comprehension prime effect of Experiment 2 and 3a, we conducted Experiment 3b as a replication of Experiment 3a with randomization.

*4.2.1. Methods*

*4.2.1.1. Participants*: We recruited an additional 360 participants for this study on Amazon Mechanical Turk. We used the same qualification criteria as in Experiments 1, 2, and 3a. As before, we excluded participants with inappropriate responses, as well as one participant who responded (presumably appropriately) in Spanish. After excluding participants who did not meet our criteria, 334 participants remained for analysis. Participants received $2 USD for their participation.

*4.2.1.2. Materials*: The counterbalance and materials for Experiment 3b are the same as for Experiment 2 and 3a.

*4.2.1.3. Procedure*: The procedure of this experiment is the same as Experiment 3b. Each question was presented on its own page. However, in this experiment, the primes and fillers were presented in a random order.

*4.2.2. Results*

*4.2.2.1. Coding*: As in Experiment 3a, to facilitate the coding process, the first author classified a random subset of all utterances into DO, PO, or Other. Then a random forest classifier was trained to predict the structure from the text and predicted labels for all remaining trials. The first author verified and corrected all of the algorithmically generated labels, with 12.2% of utterances being re-classified from the algorithmically predicted labels.

*4.2.2.2. Comprehension-to-production priming*: Participants produced either a DO or a PO 62.1% of the time in the first production phase, producing the DO form 32.6% ($SE$ = 0.03) of the time, and the PO form 29.5% of the time ($SE$ = 0.02), the opposite pattern from previous experiments (consider, for example, the 32.6% DO rate in Experiment 3b to 25.7% in Experiment 3a). In the second half of the experiment, however, speakers produced *fewer* DOs (28.4%) and *more* POs (34.1%). We aimed to construct maximal mixed effects models of the likelihood that speakers would produce a DO, but these models either failed to converge or contained correlations of 0 for some slopes. Consequently, we chose to build a model with random intercepts by participants and items only. The results of this model show that speakers produced significantly more POs (and fewer DOs) in the second half of the experiment than in the first half, the opposite pattern of results from Experiments 1, 2, and 3a, which we discuss further in the next section. These results are summarized in Table 8 below.

*4.2.2.3. Self-priming (AR1 analysis)*: We again conducted an autocorrelation regression analysis to assess self-priming. We present the statistical estimates in Table 9 below. As in Experiment 3a, the analysis found no significant self-priming effect. We present the random effects in Table 7 of Appendix B. Again, we found no evidence that speakers primed themselves in the second production phase, suggesting that comprehension input

Table 8
Mixed effects model testing for the effect of the comprehension prime on participants' likelihood of producing a Double Object (DO) structure

| Fixed Effect | Variable Type | Estimate | *SE* | z | *p* |
|---|---|---|---|---|---|
| Intercept | | −1.12 | 0.22 | −5.11 | *** |
| Before/after | Contrast coding | −0.13 | 0.04 | −3.66 | *** |

*Note*. ***indicates significance at <.001.

Table 9
Fixed effects of the Markov mixed-effect multinomial logistic regression model from Experiment 3b

| | Fixed Effects, Trial 1 | | | | | |
|---|---|---|---|---|---|---|
| | Other Versus Double Object (DO) Outcome | | | Other Versus Prepositional Object (PO) Outcome | | |
| | Estimate | l-CI | u-CI | Estimate | l-CI | u-CI |
| Grand Mean | −0.22 | −1.35 | 0.74 | −0.77 | −2.66 | 0.59 |
| AR1 for other versus DO | | – | | | – | |
| AR1 for other versus PO | | – | | | – | |
| | **Fixed Effects, Trials 2–7** | | | | | |
| | Other Versus Double Object (DO) Outcome | | | Other Versus Prepositional Object (PO) Outcome | | |
| | Estimate | l-CI | u-CI | Estimate | l-CI | u-CI |
| Mean for other | −0.20 | −1.13 | 0.53 | −0.52 | −1.40 | 0.18 |
| AR1 for other versus DO | 0.33 | −0.13 | 0.81 | 0.16 | −0.35 | 0.70 |
| AR1 for other versus PO | 0.12 | −0.39 | 0.59 | 0.52 | −0.04 | 1.11 |

*Note*. CI indicates 95% Bayesian credible interval; "l-CI" indicates the lower limit and "u-CI" indicates the upper limit. Significance marked by * for fixed effects based on 95% Bayesian credible interval testing. "–" indicates the parameters are not estimated (because they do not exist at Trial 1).

can contaminate self-priming. We include the analyses of the second production phase data in Table 8 of Appendix B.

## 5. Discussion

Experiment 3b produced similar results with respect to self-priming as Experiment 3a. Neither experiment found a significant tendency for speakers to repeat structures that they had just produced. Experiment 3a and 3b differ in that Experiment 3a replicated the comprehension-to-production effect, whereas 3b failed to find an effect of the comprehension prime in the expected direction. We discuss both of these patterns in turn below, starting with a joint analysis of the self-priming data from the first production phase of Experiments 3a and 3b.

### 5.1. Post-hoc combined self-priming analysis of Experiment 3a and 3b

The lack of a self-priming effect in both Experiment 3a and 3b is puzzling in light of the significant self-priming estimates in Experiments 1 and 2. The lack of a significant self-priming effect here may be simply due to lack of power, however, which is plausible given the effect size in the second experiment. To determine whether this might be the case, we aggregated the data from the first blocks for both Experiment 3a and 3b, for a total of 654 participants and 4,578 trials.

This joint autocorrelation analysis of Experiment 3a and 3b yielded the same results as Experiments 1 and 2, with significant self-priming effects. We present the results of these analyses below in Table 10, with the random effects in Table 9 of Appendix B. A combined analysis of the second production phase data also demonstrates a reliable self-priming effect. We present the results of this post-hoc analysis in Table 10 of Appendix B.

The results here confirm that speakers are more likely to produce a DO if they have just produced a DO, and more likely to produce a PO if they have just produced a PO. Taken together with the results of the first two experiments, we conclude that the self-priming effect can in fact be found consistently, but we caution that the effect size is highly variable, especially relative to the influence of even a single comprehension prime on production.

Two potential explanations present themselves for the lack of a self-priming effect in Experiment 3a and 3b when those experiments are analyzed individually. The first is that

Table 10
Fixed effects of the Markov mixed-effect multinomial logistic regression model combining Experiments 3a and 3b

| | Fixed Effects, Trial 1 | | | | | |
|---|---|---|---|---|---|---|
| | Other Versus Double Object (DO) Outcome | | | Other Versus Prepositional Object (PO) Outcome | | |
| | Estimate | l-CI | u-CI | Estimate | l-CI | u-CI |
| Grand Mean | −0.21 | −0.93 | 0.39 | −0.17 | −0.85 | 0.42 |
| AR1 for other versus DO | | – | | | – | |
| AR1 for other versus PO | | – | | | – | |
| | **Fixed Effects, Trials 2–7** | | | | | |
| | Other Versus Double Object (DO) Outcome | | | Other Versus Prepositional Object (PO) Outcome | | |
| | Estimate | l-CI | u-CI | Estimate | l-CI | u-CI |
| Mean for other | −0.33 | −0.79 | 0.17 | −0.35 | −0.79 | 0.03 |
| AR1 for other versus DO | 0.44 | 0.17 | 0.69* | −0.09 | −0.17 | 0.35 |
| AR1 for other versus PO | −0.14 | −0.42 | 0.12 | 0.66 | 0.41 | 0.92* |

*Note*. CI indicates 95% Bayesian credible interval; "l-CI" indicates the lower limit and "u-CI" indicates the upper limit. Significance marked by * for fixed effects based on 95% Bayesian credible interval testing. "–" indicates the parameters are not estimated (because they do not exist at Trial 1).

Experiments 1 and 2, by virtue of presenting all of the items simultaneously, encouraged participants to revise prior descriptions or refer to prior descriptions when making later syntactic choices. A second possibility is that participants may have copied and pasted their responses to save time. In this case, apparent self-priming in Experiments 1 and 2 could be due at least in part to task demands. If task demands explain the bulk of self-priming, then it is difficult to integrate these results with either error-driven learning accounts of syntactic priming or activation-based accounts, as selection would no longer be a (presumably) unconscious, implicit decision.

We think the above explanations are unlikely. It is doubtful that participants revised earlier answers after completing later answers, given the amount of time this would have taken. Individuals who took longer than 25 min for any of the experiments were excluded from the analysis. In addition, we know that workers on Mechanical Turk are generally paid significantly below the American minimum wage on average (Hara et al., 2018). Thus, participants are incentivized to complete studies as quickly as possible, so that they can move on to the next paid task. This is supported by the amount of time participants took to complete the task. Participants took approximately 12–14 min on average to answer a total of 50 questions (34 of them production trials). Thus, although we cannot rule out this revision explanation, we think it is unlikely. The other possibility—that participants reused responses—is also unlikely. If this were the case, then we would expect much greater syntactic reuse than we observe. Individuals' biases are relatively stable with very few participants producing only one of a single syntactic structure. Moreover, copying and revising pasted responses may take more time than writing the responses from scratch.

We think that it is more likely that self-priming effects are simply just small, which is supported by the size of the effects we saw in Experiments 1 and 2. The fact that self-priming re-emerges in a combined analysis of Experiment 3a and 3b suggests that in a task more similar to classic experiments of syntactic priming, speakers can still prime themselves.

### *5.2. Comprehension-to-production priming in 3a versus 3b*

A puzzle still remains: Why does DO priming lead to *fewer* DOs in Experiment 3b when DO priming led to more DOs produced in every other experiment in this paper? There are a number of possible explanations, but the most obvious difference between the experiments before considering other factors is that speakers were more likely to produce a DO than a PO at every single time point in Experiment 3b. The data from Experiment 3b are anomalous in that DO and PO appear relatively equibiased, whereas in all other experiments in the present study, DO was the minority structure. This may be evidence of a ceiling effect, which would limit our ability to detect comprehension-to-production priming for the target structure of interest. Similarly, what we may be observing for the increase in PO usage might simply reflect regression to the mean. Which of these is more likely is a topic of future research. Critically, we *do* find evidence for comprehension-to-production priming in three of the four experiments presented here, and this effect is well attested in the literature.

## 6. General discussion

In all three experiments, we find that speakers are more likely to produce a construction if they have just recently produced that construction themselves. The evidence for self-priming presented here is consistent with a number of corpus studies with similar findings (e.g., Gries, 2005; Jaeger & Snider, 2013; Myslín & Levy, 2016; Reitter et al., 2011). The current study is novel in that we see evidence for self-priming while ruling out alternative explanations. Previous studies examined structural repetition in corpora of natural speech, in which other factors such as discourse structure, speaker biases, or task demands could be the driving force behind self-priming effects. Furthermore, experimental studies that have demonstrated self-priming have largely employed paradigms restricting speakers' choice in syntactic selection (Kaschak et al., 2014; Sagaert et al., 2013). In the current study, we control for these factors and still find evidence of self-priming, though the effect size is small and may not be easily observed in laboratory studies with relatively small numbers of participants and items.

Comprehension-to-production effects are more robust, demonstrating that self-priming and comprehension input differ in their detectability. A speaker's linguistic choices in production are highly sensitive to what that speaker has recently heard. In Experiment 1, 2, and 3a, speakers were more likely to produce a construction in the second production phase after hearing that construction in the comprehension phase, replicating a vast number of previous findings (see Mahowald et al., 2016 for a comprehensive review). The effect is robust even when speakers process a single instance of a grammatical structure —in Experiments 2 and 3a, we found evidence that comprehension-to-production priming can occur after only a single exposure to a construction in the comprehension phase. We failed to find comprehension-to-production priming in Experiment 3b, where participants actually produced fewer instances of the prime structure after exposure, which may have been due to participants' overall preference for the DO structure, or may have simply been an anomalous result .

Finally, we also replicated the finding that the magnitude of the comprehension-to-production priming effect for a specific syntactic structure is directly related to the probability of that syntactic structure, which is also consistent with previous work on the inverse frequency effect (Bock, 1986; Ferreira, 2003; Hartsuiker & Westenberg, 2000; Mahowald et al., 2016).

The three effects of self-priming, comprehension-to-production priming, and the inverse frequency effect shed light on activation-based models and error-driven models of syntactic structure, which we review again below.

### *6.1. Activation- and exemplar-based models best account for self-priming*

In residual activation models, syntactic structures are selected and retrieved through spreading activation. The retrieval of a given syntactic structure is strengthened by multiple retrievals and productions (Pickering & Branigan, 1999; Reitter et al., 2011). Because of this increased activation, speakers are more likely to continue to produce a structure that they have just said relative to syntactic alternatives. The same is true for comprehension—if a listener hears a particular structure from another speaker, the activation for this structure is strengthened, which increases the likelihood that the structure will be selected for production in the future. Error-driven models typically do not predict self-priming without an error signal. In these frameworks, syntactic selection is a zero-sum game. Priming occurs when listeners make predictions about upcoming syntactic structures. If a syntactic structure is encountered that violates expectations, an error signal is generated that leads to an adjustment of future expectations, and listeners are now more likely to expect the unexpected structure (Chang et al., 2006). Explanations of comprehension-to-production priming are straightforward under an error-based account: Speakers are more likely to produce a syntactic structure that they have just heard, because that recently encountered structure increases the *relative* frequency of the two structures in the speaker's internal model. However, most error-driven models cannot explain self-priming. Because a speaker's own productions cannot generate an error signal, they cannot affect a speaker's expectations for upcoming structures. If this is the case, the self-priming results here are inconsistent with error-based models. Unlike error-based models, both activation-based (Reitter et al., 2011) and exemplar-based models of syntactic priming such as those of Jaeger and Snider (2013) simultaneously show sensitivity to comprehended structures *and* account for self-priming effects in production.

### 6.2. Error-based models best account for inverse frequency effects

However, one advantage of error-driven learning models over activation models is that they explicitly capture the inverse frequency effect, in which low-probability structures have a greater influence on speakers' behaviors than high-probability structures. We see this in our data, where comprehending an utterance containing the lower-probability structure (DO, in these stimuli) had a larger influence on production than comprehending higher-probability structures (PO; Experiment 1). Activation-based models such as Reitter et al. (2011) account for the inverse frequency effect only incidentally—the weights associated with less active structures show greater *relative* change, but low-probability structures do not directly affect the size of priming as with error-based models. The differential influence of DO primes and PO primes on production is difficult to reconcile with activation accounts.

The larger priming effect for DO primes (the dispreferred structure overall) over PO primes fits neatly, however, with error-driven learning accounts (e.g., Chang et al., 2006; Jaeger & Snider, 2013). In these accounts, listeners make specific predictions about upcoming syntactic structures using their own internal models of the relative frequency of syntactic constructions and adjust their behavior when they are wrong—when the structure that was processed (either as input via comprehension or the model's own productions) matches the intended message, there is little error. The size of this error determines the size of priming, with more unusual structures leading to greater priming.

### 6.3. Alternate models could support self-priming and comprehension-to-production priming

We may be able to gain insight into how models of syntactic priming work by examining models of adaptation processes in other domains of language production, such as lexical selection (Howard, Nickels, Coltheart, & Cole-Virtue, 2006; Mahon, Costa, Peterson, Vargas, & Caramazza, 2007; Oppenheim, Dell, & Schwartz, 2010). For example, Oppenheim et al. (2010) proposed an activation-based account of lexical production. In their model, repetition of a word (e.g., *cow*) reduces activation of related (neighboring) words (e.g., *horse*, *sheep*, or *goat*). The result is that the activation of the most recently produced word becomes stronger, and semantically related words lose activation. A similar computation mechanism may play a role in syntactic production. In the same way that lexical items can compete with each other for activation, syntactic structures are thought to compete with each other for selection (i.e., Bock, 1986).

While there is little research in this vein suggesting that syntactic selection operates in the same way as lexical selection processes, "competition-based" syntactic selection mechanisms would also make the interesting prediction that speakers can learn from their own productions. Because competition-based models allocate more activation to recently produced words or constructions (e.g., a DO) by reducing activation of other constructions (e.g., a PO), this leads to a relative dispreference for selecting the alternate constructions. We note that competition-based accounts, which lack an explicit error mechanism, are conceptually similar to exemplar-based Bayesian belief updating models of syntactic alternations (e.g., Kleinschmidt, Fine, & Jaeger, 2012). Neither framework requires the calculation of error, simply the updating of weights (or probabilities) to increase the activation of recently produced constructions (e.g., a DO), which must be done by taking activation from the other construction. Whether competition-based models or exemplar-based models are a better fit to the behavioral data is an interesting empirical question that merits further research.

### 6.4. Learning and memory mechanisms involved in self-priming and comprehension-to-production priming may differ

All theories of syntactic priming predict long-term structural persistence following exposure to a specific syntactic structure; experience permanently changes the production system (Bock & Griffin, 2000; Chang et al., 2006; Dell & Chang, 2014; Jaeger & Snider, 2013; Reitter et al., 2011). We see clear, long-lasting effects of syntactic comprehension on production that persist throughout the second production phase of our experiments. Experiment 3a provides some evidence that these results cannot be attributed entirely to lexical repetition, because speakers' syntactic preferences changed after exposure to a single prime presented at the very beginning of the comprehension phase, which would have been several minutes before the second production phase on average, well outside the range of lexically mediated priming effects (Hartsuiker & Kolk, 1998; Hartsuiker et al., 2008; Tooley & Traxler, 2018).

The results are less clear for self-priming. The autocorrelation analyses confirmed the presence of temporal dependencies between speakers' productions. However, we do not know whether self-priming has long-lasting effects on the production system, or whether self-priming is limited to the very near-term. If self-priming is due to lexical repetition (a potential problem identified by Healey et al., 2014 and a critical component of Reitter et al., 2011) or does not lead to longer-term structural priming, then we might conclude that self-priming is accomplished by different selection and learning mechanisms within the production system than comprehension-to-production priming. Identifying how similar the learning and memory mechanisms are between the two processes will be an important next step in building a comprehensive model of syntactic priming that hopes to account for speakers' decisions in the absence of exposure to another person's speech.

Thus, we find some support for both activation-based and error-based theories of syntactic priming. While all models predict comprehension-to-production priming, activation models are most consistent with self-priming, and error-driven models most consistent with the inverse frequency effect. In the absence of evidence, however, we propose that the best account of the effects we find here, as well as effects in the broader literature, point towards a hybrid model that integrates activation and error-driven learning, perhaps similar in spirit to Reitter et al. (2011), Jaeger and Snider (2013), or Tooley and Traxler (2018)'s account of transient and long-lasting syntactic priming in comprehension.

The small comprehension-to-production priming effects for a single prime in Experiments 2 and 3 show that the presence of a comprehension-to-production priming effect is not always as influential as self-priming effects on language production. In contrast to the failure to find comprehension-to-production priming in Experiment 3b, we found small self-priming effects in all populations, though it is clear that there are still substantial individual differences that make it difficult to estimate self-priming in small samples. The fact that it is possible to find evidence of self-priming but not comprehension-to-production priming may constitute further evidence that the learning mechanisms involved in self-priming are at least partially distinct.

Finally, the fact that we see evidence for self-priming raises an interesting question: The logical end-state of a system that can prime itself would be to eventually choose a syntactic structure at the exclusion of all other alternatives. Why doesn't this happen? We can only speculate that outside of the laboratory, a speaker's linguistic choices are largely driven by the message they mean to convey, pragmatic and discourse constraints, social constraints, and other cognitive factors, all of which conspire to create variability in the linguistic signal, which keeps us from sounding like a syntactically broken record.

In sum, by using a well-known statistical technique in a novel way, we have presented evidence that speakers prime themselves syntactically while also replicating previous comprehension-to-production priming effects and the inverse frequency effect. Altogether, our results suggest that syntactic priming depends both on error-driven learning and on residual activation.

## Notes

1. https://osf.io/tcskf/
2. Though we do not present the analyses here, we ran a post hoc analysis of PO usage using the same analysis as in Table 4. This analysis replicates the results from the DO condition in Experiment 1, with a significant drop in PO usage between blocks.

## References


Babcock, M. J., Ta, V. P., & Ickes, W. (2014). Latent semantic similarity and language style matching in initial dyadic interactions. *Journal of Language and Social Psychology*, *33*, 78–88.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. *R Package Version*, *1*, 1–23.

Bernolet, S., & Hartsuiker, R. J. (2010). Does verb bias modulate syntactic priming? *Cognition*, *114*, 455–461.

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

Bock, K., Dell, G. S., Chang, F., & Onishi, K. H. (2007). Persistent structural priming from language comprehension to language production. *Cognition*, *104*, 437–458.

Bock, K., & Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology: General*, *129*, 177–192.

Bock, K., & Loebell, H. (1990). Framing sentences. *Cognition*, *35*, 1–39.

Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis: Forecasting and control* (Revised ed). San Francisco, CA: Holden-Day.

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (1999). Syntactic priming in written production: Evidence for rapid decay. *Psychonomic Bulletin & Review*, *6*, 635–640.

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, *113*, 234–272.

Cho, S.-J., Brown-Schmidt, S., & Lee, W.-Y. (2018). Autoregressive generalized linear mixed effect models with crossed random effects: An application to intensive binary time series eye-tracking data. *Psychometrika*, *83*, 751–771.

Dell, G. S., & Chang, F. (2014). The P-chain: relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*, 20120394. https://doi.org/10.1098/rstb.2012.0394.

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*, 379–398.

Fine, A. B., & Jaeger, T. F. (2013). Evidence for implicit learning in syntactic comprehension. *Cognitive Science*, *37*, 578–591.

Gries, S. T. (2005). Syntactic priming: A corpus-based approach. *Journal of Psycholinguistic Research*, *34*, 365–399.

Hara, K., Adams, A., Milland, K., Savage, S., Callison-Burch, C., & Bigham, J. P. (2018). *A data-driven analysis of workers' earnings on Amazon Mechanical Turk. In Proceedings of SIGCHI Conference on Human Factors in Computing Systems, CHI '18*. New York: ACM.

Hartsuiker, R. J., Bernolet, S., Schoonbaert, S., Speybroeck, S., & Vanderelst, D. (2008). Syntactic priming persists while the lexical boost decays: Evidence from written and spoken dialogue. *Journal of Memory and Language*, *58*, 214–238.

Hartsuiker, R. J., & Kolk, H. H. (1998). Syntactic persistence in Dutch. *Language and Speech*, *41*, 143–184.

Hartsuiker, R. J., & Westenberg, C. (2000). Word order priming in written and spoken sentence production. *Cognition*, *75*, B27–B39.

Healey, P. G., Purver, M., & Howes, C. (2014). Divergence in dialogue. *PLoS ONE*, *9*, e98598.

Hedeker, D. (2003). A mixed-effects multinomial logistic regression model. *Statistics in Medicine*, *22*, 1433–1446.

Howard, D., Nickels, L., Coltheart, M., & Cole-Virtue, J. (2006). Cumulative semantic inhibition in picture naming: Experimental and computational studies. *Cognition*, *100*, 464–482.

Ireland, M. E., Slatcher, R. B., Eastwick, P. W., Scissors, L. E., Finkel, E. J., & Pennebaker, J. W. (2011). Language style matching predicts relationship initiation and stability. *Psychological Science*, *22*, 39–44.

Jaeger, T. F., & Snider, N. E. (2013). Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition*, *127*, 57–83.

Kaschak, M. P. (2007). Long-term structural priming affects subsequent patterns of language production. *Memory & Cognition*, *35*, 925–937.

Kaschak, M. P., Kutta, T. J., & Coyle, J. M. (2014). Long and short term cumulative structural priming effects. *Language, Cognition and Neuroscience*, *29*, 728–743.

Kaschak, M. P., Kutta, T. J., & Schatschneider, C. (2011). Long-term cumulative structural priming persists for (at least) one week. *Memory & Cognition*, *39*, 381–388.

Kleinschmidt, D. F., Fine, A. B., & Jaeger, T. F. (2012). A belief-updating model of adaptation and cue combination in syntactic comprehension. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th annual conference of the cognitive science society* (pp. 599–604). Austin, TX: Cognitive Science Society.

Mahon, B. Z., Costa, A., Peterson, R., Vargas, K. A., & Caramazza, A. (2007). Lexical selection is not by competition: A reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 503.

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.

Morey, R. D., Hoekstra, R., Rouder, J. N., Lee, M. D., & Wagenmakers, E. J. (2016). The fallacy of placing confidence in confidence intervals. *Psychonomic Bulletin & Review*, *23*, 103–123.

Myslín, M., & Levy, R. (2016). Comprehension priming as rational expectation for repetition: Evidence from syntactic processing. *Cognition*, *147*, 29–56.

Oppenheim, G. M., Dell, G. S., & Schwartz, M. F. (2010). The dark side of incremental learning: A model of cumulative semantic interference during lexical access in speech production. *Cognition*, *114*, 227–252.

Pettitt, A., Tran, T., Haynes, M., & Hay, J. (2006). A Bayesian hierarchical model for categorical longitudinal data from a social survey of immigrants. *Journal of Royal Statistical Society-Series A*, *169*, 97–114.

Pickering, M. J., & Branigan, H. P. (1999). Syntactic priming in language production. *Trends in Cognitive Sciences*, *3*, 136–141.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*, 169–190.

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, *36*, 329–347.

Reitter, D., Keller, F., & Moore, J. D. (2011). A computational cognitive model of syntactic priming. *Cognitive Science*, *35*, 587–637.

Reitter, D., & Moore, J. D. (2014). Alignment and task success in spoken dialogue. *Journal of Memory and Language*, *76*, 29–46.

Seedorff, M., Oleson, J., & McMurray, B. (2018). Detecting when timeseries differ: Using the Bootstrapped Differences of Timeseries (BDOTS) to analyze Visual World Paradigm data (and more). *Journal of Memory and Language*, *102*, 55–67.

Segaert, K., Kempen, G., Petersson, K. M., & Hagoort, P. (2013). Syntactic priming and the lexical boost effect during sentence production and sentence comprehension: An fMRI study. *Brain and Language*, *124*, 174–183.

Skrondal, A., & Rabe-Hesketh, S. (2003). Multilevel logistic regression for polytomous data and rankings. *Psychometrika*, *68*, 267–287.

Spiegelhalter, D., Thomas, A., & Best, N. (2003). *WinBUGS version 1.4.3 [Computer program]*. Cambridge, UK: MRC Biostatistics Unit, Institute of Public Health.

Szmrecsanyi, B. (2008). *Morphosyntactic persistence in spoken English. A corpus study at the intersection of variationist sociolinguistics, psycholinguistics, and discourse analysis*. Berlin: De Gruyter Mouton.

Tooley, K. M., & Bock, K. (2014). On the parity of structural persistence in language production sand comprehension. *Cognition*, *132*, 101–136.

Tooley, K. M., & Traxler, M. J. (2018). Implicit learning of structure occurs in parallel with lexically-mediated syntactic priming effects in sentence comprehension. *Journal of Memory and Language*, *98*, 59–76.

## Supporting Information

Additional Supporting Information may be found online in the supporting information tab for this article:

**Appendix A.** Stimuli.

**Appendix B.** Supplemental analyses.

**Appendix C.** A description of a multivariate generalized linear mixed effect model.