Cognition 197 (2020) 104183

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Syntactic category constrains lexical competition in speaking

Shota Momma[a,*], Julia Buffinton[b], L. Robert Slevc[b], Colin Phillips[b]

[a] University of Massachusetts, Amherst, United States of America
[b] University of Maryland, College Park, United States of America

ARTICLE INFO

Keywords:
Lexical access
Syntactic category
Grammatical encoding
Sentence production
Sentence picture interference

ABSTRACT

We report two experiments that suggest that syntactic category plays a key role in limiting competition in lexical access in speaking. We introduce a novel sentence-picture interference (SPI) paradigm, and we show that nouns (e.g., *running* as a noun) do not compete with verbs (e.g., *walking* as a verb) and verbs do not compete with nouns in sentence production, regardless of their conceptual similarity. Based on this finding, we argue that lexical competition in production is limited by syntactic category. We also suggest that even complex words containing category-changing derivational morphology can be stored and accessed together with their final syntactic category information. We discuss the potential underlying mechanism and how it may enable us to speak relatively fluently.

1. Introduction

Saying a word requires accessing an appropriate representation of the word among tens of thousands of words in speakers' mental dictionaries, many of which are similar to each other. Lexical access requires overcoming competition from these similar words, and competition is likely even greater when saying a sentence because speakers must rapidly access multiple words in a specifically ordered sequence, while each accessed word creates an additional source of interference for the others. Yet healthy adult native speakers produce sentences mostly fluently and relatively effortlessly. In this paper, we report two experiments that suggest that syntactic category plays a key role in limiting competition in lexical access in speaking. We introduce a novel sentence-picture interference (SPI) paradigm, and we show that nouns (e.g., *running* as a noun) do not compete with verbs (e.g., *walking* as a verb) and verbs do not compete with nouns in sentence production, regardless of their conceptual similarity. Based on this finding, we argue that lexical competition in production is limited by syntactic category and that even complex words containing category-changing derivational morphology can be stored and accessed together with their syntactic category information.

1.1. Competition in lexical access

In single word production research, it is widely assumed that items in lexical memory are selected competitively (Levelt, Roelofs, & Meyer, 1999; Roelofs, 1992). By "items in lexical memory," we specifically mean *lemmas* (Kempen & Huijbers, 1983), which are abstract linguistic representations that contain syntactic and semantic information but not phonological information (Levelt et al., 1999). To select a lemma involves resolving competition from non-target lemmas. This interference is especially strong from conceptually similar competitors that belong to the same semantic category (i.e., *co-hyponyms* of the target words). For instance, when accessing the lemma for *cat*, activation of a conceptually similar lemma such as *dog* can interfere with target retrieval, causing delays and/or increased errors in production. This pattern is often observed in experimental settings such as the picture-word interference (PWI) task where speakers name pictures while ignoring distractor words that are presented together or temporally closely. Conceptually similar distractor words delay articulation onset of picture-name production compared to unrelated distractors (Lupker, 1979; Roelofs, 1992; Schriefers, Meyer, & Levelt, 1990; Vigliocco, Vinson, & Siri, 2005).

Investigations of this semantic interference effect have usually been limited to a specific kind of conceptual relation, specifically the relation between the members of the same category (e.g., *dog* and *cat*). The magnitude of this effect is proportional to the conceptual similarity between the target and the non-target words (Vigliocco, Vinson, Lewis, & Garrett, 2004). The target picture is assumed to activate a cat concept which spreads activation both to the target lemma (e.g., *cat*) and to conceptually related non-target lemmas (e.g., *dog*). When non-target lemmas receive extra activation from distractors, the contrast in activation becomes low, causing delays in production and/or speech errors.

When producing a sentence, an even larger set of lemmas may be

* Corresponding author at: Department of Linguistics, University of Massachusetts, Amherst, United States of America.
E-mail address: snegishi@umass.edu (S. Momma).

https://doi.org/10.1016/j.cognition.2020.104183
Received 6 June 2018; Received in revised form 27 December 2019; Accepted 2 January 2020
0010-0277/

co-active during each retrieval process, as speakers need to retrieve multiple lemmas in a rapid sequence. This seemingly poses a challenge for speakers if these multiple words compete with each other. Indeed, intra-sentence competition can result in speech errors, especially whole word exchange errors such as erroneously saying *the frisbee is catching a dog* when intending to say *the dog is catching a frisbee* (Fromkin, 1971; Garrett, 1975). This type of error is commonplace in naturalistic speech. Detailed observations of word exchange error patterns in speech corpora have revealed that this type of error is subject to a syntactic category constraint (Fromkin, 1971; Garrett, 1975, 1980; Nooteboom, 1980).

Though sometimes considered not absolute (Stemberger, 1985), this constraint prevents exchange errors between two words that differ in their syntactic category such that nouns only exchange with nouns, and verbs only with verbs. Under the assumption that exchange errors reflect a failure to resolve competition properly, this speech error evidence suggests that words that differ in syntactic category do not compete with each other (but see Discussion for alternative accounts of the syntactic category constraint). Dell, Oppenheim, and Kittredge (2008) implemented a hypothetical mechanism for this constraint, charmingly named the *syntactic traffic cop*, in their connectionist model of word production. In their model, every lemma is connected to syntactic category nodes (representing noun, verb, etc.) with appropriate connection weights, and the sentence context activates syntactic category nodes differentially. For example, noun nodes activate all nouns and inhibit the activation of all verbs. This limits competition between words of different categories. The role of the sentence context in Dell and colleagues' model fits with evidence from PWI experiments that pit conceptual similarity against syntactic category. Specifically, Vigliocco et al. (2005) showed that, in Italian, saying a verb (e.g., *ridere*; roughly *to laugh*) is delayed more by conceptually related distractor verbs (e.g., *sospirare*: roughly *to sigh*) than by conceptually related distractor nouns (e.g., *il pianto*; roughly *the cry*) when semantic similarity is closely matched (according to their previous model that quantifies the similarity between words; Vigliocco et al., 2004). This additional interference from syntactic category overlap was only observed when speakers produced inflected verbs, and not when speakers produced the uninflected citation form. They interpreted this to mean that syntactic category overlap creates an additional source of interference, but only when speakers produce a syntactically complex utterance. Dell et al. (2008) argued that this additional interference effect from shared syntactic category is explained by their syntactic traffic cop mechanism.

Understanding the syntactic category constraint is critical in bridging the existing gap between word production and sentence production models, as it is a constraint that arises from currently unknown interactions between sentential and lexical processes. However, the mechanism underlying the syntactic category constraint is not well understood. Indeed, the syntactic category constraint on speech errors does not have to be characterized as a constraint on lexical selection. For example, it can be characterized as a constraint on post-selection processes, such as integration of the selected words to overall sentence structure (see General Discussion for more detailed discussion). Previous experimental studies, which could potentially help us understand the underlying mechanism more, do not neatly align with the speech error data. For example, the study by Vigliocco and colleagues described above was used as convergent evidence for the view that syntactic category directly constrains lexical retrieval. However, Vigliocco and colleagues did find cross-category interference effects. What they showed was an additional effect of syntactic category overlap, so their data do not provide evidence that the syntactic category has a competition limiting function, or more strongly, competition eliminating function. Word exchange and substitution errors that violate the syntactic category constraint are very rare, so it is unclear why speakers experienced reliable interference from conceptually similar noun distractors (a 27 ms effect). Furthermore, note that this cross-category interference effect was not especially small for a PWI paradigm. For instance, here are the magnitude of semantic interference effects observed in well-known PWI-studies: Lupker (1979) found a 32 ms effect, Schriefers et al. (1990) found a 40 ms interference effect, and Damian and Bowers (2003) found a 16 ms interference effect. Based on the syntactic category constraint observed in the speech error record, speakers should be expected to experience little or no interference from conceptually similar distractors from a different syntactic category. Other work relevant to the category constraint comes from Pechmann and colleagues (Pechmann, Garrett, & Zerbst, 2004; Pechmann & Zerbst, 2002), who showed that distractors from the same syntactic category induce an interference effect in the absence of semantic relatedness. However, these results are disputed. Janssen, Melinger, Mahon, Finkbeiner, and Caramazza (2010) showed that the greater interference from the same-category distractors (i.e., content words) were due to the higher imageability of content words compared to the across-category distractors (i.e., function words). In any case, Pechmann and colleagues' results do not show that syntactic category has a competition-limiting or competition-eliminating function, because they simply show that there is additional interference from syntactic category overlap.

What, then, explains the disconnect between the speech error record and the cross-category interference effect observed by Vigliocco et al. (2005)? One possibility is that the noun distractors in Vigliocco et al.'s experiments might not have been perceived unambiguously as nouns. This is because the noun distractors in this PWI experiment were determiner noun sequences in Italian (e.g., *il pianto*; *the cry*). However, event denoting nouns in Italian are often homophonous with some forms of root verbs (e.g., the past-participle form of *piangere* is homophonous with the nominal *pianto*), as was the case in the majority of distractors used by Vigliocco and colleagues. Thus the noun distractors in this study, like *il pianto*, might initially have been processed as verbs (e.g., *pianto*; *cried*) before being constrained to a nominal form via combination with *il*.

A second possibility is that the manipulation of syntactic category is the study by Vigliocco et al. (2005) was confounded with morphological overlap. Specifically, Italian has three classes of verbs that end with *-are*, *-ere*, or *-ire* in their infinitive form. These three classes of verb show different morphophonological patterns in a 3rd-person singular context and other contexts. If a morphophonological pattern mismatch between the distractor and target verbs causes delays in conjugating a verb prior to production, then the observed interference effect cannot unambiguously be attributed to the overlap in syntactic category. Indeed, we counted the number of verb class mismatches in Vigliocco and colleagues' stimuli, and found that their stimuli contained 26 target-distractor pairs that mismatched in morphological class (out of 40 pairs in the verb conditions, equally distributed between the semantically related and unrelated conditions). This may explain why they found interference from syntactic category overlap only when speakers uttered inflected verbs and not when they uttered uninflected verbs in their control experiment.

Because of these concerns with previous experimental work, it remains unclear whether syntactic category does, in fact, limit competition in speaking as is suggested from the speech error record. Thus, the first aim of the current article is to re-evaluate whether syntactic category limits interference in lexical access, and if so, how strongly.

1.2. Current experiments

The current study tests whether syntactic category limits competition in lexical access during sentence production by examining patterns of interference in a modified PWI paradigm. To test these hypotheses, there are some experimental challenges that need to be overcome. The above discussion of Vigliocco et al. (2005) and Pechmann and Zerbst (2002), Pechmann et al. (2004) reveals how hard it is to use the PWI paradigm to manipulate the syntactic category of distractors. The problems arise mainly from the fact that the syntactic category of a word is

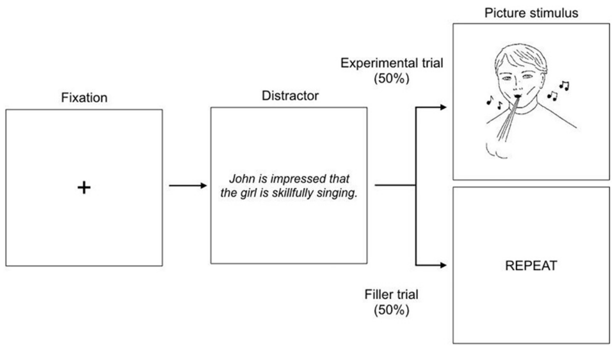

Fig. 1. Graphical illustration of the basic task structure in the Sentence-Picture Interference task.

often ambiguous without proper morphological and syntactic context, and the PWI task is unsuited to supplying context to distractors without introducing confounds or without violating the fundamental assumption of PWI, that the relevant property of the distractor (in this case, syntactic category) must be extractable automatically and rapidly.

To overcome these challenges, we introduce a novel experimental task that we named the sentence-picture interference (SPI) task, illustrated in Fig. 1. This task is similar to PWI, but with two critical differences. First, in SPI, both the distractor and the target are sentences rather than single words or phrases. This makes it possible to supply morphological and syntactic context to the distractor words while minimizing conceptual and phonological confounds, as in the following sentence pair.

(1) *John is impressed that the girl is skillfully singing.* [Verbal context].
(2) *John is impressed by the girl's skillful singing.* [Nominal context].

In this sentence pair, the underlined word *singing* is verbal in (1) and nominal in (2). Phonologically, the critical words *singing* are identical. Conceptually, the verb *singing* and the noun *singing* are maximally similar to each other. Of course, the conceptual meanings might not be identical, but the relevant question is whether such a difference can modulate semantic interference effects. We consider it unlikely that any subtle conceptual difference between the verbal and nominal versions of *singing* should be larger than the conceptual differences between associated items that routinely elicit semantic interference effects, such as *cat* and *dog*.

Second, in SPI (unlike PWI), speakers do not ignore the distractor but instead are asked to memorize the distractor sentences and are tested for their memory on half of the trials. The basic task structure is illustrated in Fig. 1. This memory test ensures that the distractor sentences are active in speakers' minds (and thus potentially able to cause interference) even on non-test trials when they produce target sentences in response to picture stimuli. There are multiple possible ways in which speakers may retain the distractor sentence representations in memory; for example, speakers may verbally rehearse the memorized sentences (using the phonological loop, Baddeley & Hitch, 1974), or they may simply retain the meaning behind sentences (Potter & Lombardi, 1990, Sachs, 1967). But, regardless of how speakers choose to encode distractor sentences in memory, the activation of relevant distractor lemmas should be heightened and should not be suppressed until the distractor sentence becomes irrelevant to speakers' task. We thus hypothesize that this memory requirement in the SPI task makes the lemmas in distractor sentences potential competitors for subsequent production.

Using the SPI task, we tested whether complex nominal gerunds like *singing* interfere with conceptually similar progressive verbs like *whistling*. If they do, it suggests either that syntactic category does not limit retrieval interference, or that the target of the retrieval process is not specified for its syntactic category. More interestingly, if *singing* as a noun and *whistling* as a verb do not interfere with each other, syntactic category needs to be able to limit retrieval interference, suggesting that the targets of retrieval are specified for their syntactic category. Thus, the current study is informative both about the role of syntactic category, and also the nature of our lexical memory used in speaking. The two experiments that we report below use the SPI task, and they both test the critical question of the current article: whether words from two different syntactic categories compete with each other.

Before describing the details of Experiments 1 and 2, it is worth discussing our rationale for the design of the SPI task in more depth, especially given that the SPI is a new task and given a potential concern that this task might at first sight seem too complex. First, the SPI requires that speakers memorize the distractor sentences. This is because the distractor words need to be kept active in order to exert an effect on the subsequent target production. Without this memorization requirement, it is not clear if the activation-level of distractors should be high enough to cause interference in subsequent target production. Relatedly, the SPI task requires task switching between sentence recall and picture description. This is in order to ensure that speakers actually maintain the distractor sentences in their memory. Furthermore, given the current theoretical question of whether cross-category words interfere with each other, the SPI uses sentence distractors rather than word distractors because of the fact that syntactic category is determined by the syntactic and morphological contexts. Similarly, given the current theoretical question, the SPI uses a sentence production task rather than a word production task, because single-word production does not allow clear disambiguation of the syntactic category of the target words. Thus, each task component of the SPI task is justified, and its design is a natural consequence of addressing the requirements for unambiguously manipulating the syntactic categories of distractor and target words without changing semantic and phonological aspects of them.

Fig. 2 shows a schematic illustration of the distractor and target activation as a function of time. This illustrative model abstracts away from many details and only captures the relevant aspects of the tasks. But, critically, it illustrates our hypothesis that, due to the memorization requirement of the SPI task, the activation of distractor words resists decay until the target picture appears. As a result, the time windows for distractor activation and target activation overlap. When the target picture is related to the distractor word, the suppression or decay of the activation level of the distractor word can be delayed due to the partial semantic overlap between the concept depicted in the picture and the distractor word, resulting in a semantic interference effect just like in PWI tasks.

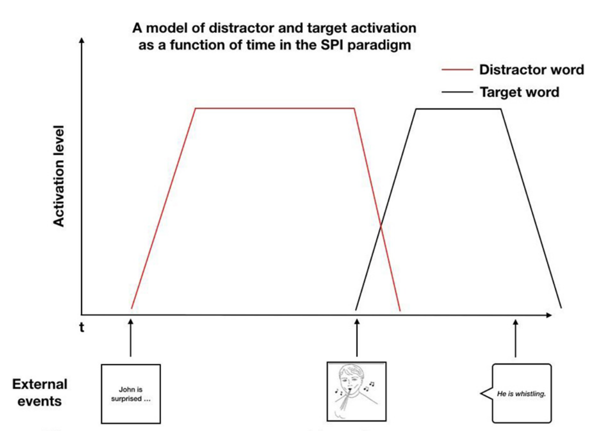

Fig. 2. Graphical illustration of our hypothesis about the time-course of distractor and target activation in the SPI task.

1.2.1. Testing the SPI task

Because the SPI task is a new task that has not been used before, we first conducted a simple two-condition experiment that tested whether a semantic interference effect can be obtained using the SPI task. In this experiment, participants ($n = 24$) memorized a sentence ending with a distractor word, and subsequently (in 50% of the trials, just like in Experiment 1 and 2) named a simple object picture. This experiment showed that a semantic interference effect from the last word of distractor sentences can indeed be obtained, just like in PWI experiments. Thus, we are justified to use SPI to investigate the question of whether two words from different syntactic categories interfere with each other. The design and results of this experiment are reported in the Appendix.

2. Experiment 1

Experiment 1 tested whether verbs interfere only with verbs and not with nouns. We used the SPI paradigm for this purpose. This is a critical test for the hypothesis that syntactic category limits lexical competition. In Experiment 1, we measured verb production latency rather than speech onset latency, because it has been demonstrated multiple times that semantic interference effects can occur later than the sentence onset (Momma, Slevc, & Phillips, 2018), so using the sentence onset latency measure can be misleading because potential semantic interference effects in the speech duration measures may be missed.

2.1. Method

2.1.1. Participants

Forty-eight University of Maryland undergraduate students participated in Experiment 1 for course credit. Informed consent was obtained from all participants prior to the experiment. Three participants were replaced from the analysis for not following instructions, and one participant was replaced because English was not his or her native language.

2.1.2. Materials

Twenty-four pictures of actions corresponding to (optionally) intransitive verbs (e.g., *sing*, *cook*, *whistle*, *run*, *walk*, *cook*, etc.) were selected from the UCSD International Picture Naming Database (Szekely et al., 2004). Forty-eight distractor sentences containing a word that was semantically related to the critical target word were constructed and paired with the target picture to create semantically related target-distractor pairs. Semantic relatedness was assessed based on intuitive judgments and verified using cosine distance in Latent Semantic Analysis (Landauer & Dumais, 1997). The related distractor sentences were each re-paired with another target picture to create unrelated target-distractor pairs. That is, the sets of related and unrelated distractor sentences were identical. The mean cosine distance between the target and the last word of the distractor sentence was 0.42 (sd = 0.11) in the related pairs and 0.12 (sd = 0.08) in the unrelated pairs (two-tailed *t*-test; $t(46) = -10.84$, $p < .001$). An example picture stimulus and example distractor sentences are shown in Fig. 3. Each participant saw

Distractor relatedness	Distractor category	Distractor sentence	Target picture
Related	Verbal	John is impressed that the girl is skillfully singing.	
Related	Nominal	John is impressed by the girl's skillful singing.	
Unrelated	Verbal	Mary told the doctor that she is persistently coughing.	
Unrelated	Nominal	Mary told the doctor about her persistent coughing.	

Fig. 3. Example distractor sentences used in Experiment 1. Note that all unrelated distractor sentences were also used as related distractor sentences in other trials for a different participant, i.e., the sentential frame *John is impressed that/by* was used in unrelated conditions as well, and the sentential frame *Mary told the doctor that/about* was used in related conditions as well.

each sentence distractor sentence only once. Also, in some of the distractor sentences, the critical word was not necessarily the last word of the sentence (though it was the last word in 42 of 48 distractor sentences). The complete list of target verbs and distractor sentences used in Experiment 1 (and Experiment 2) is available at https://shotam.github.io/CategorySPI/StimList.csv.

2.1.3. Procedure and analysis

Each experimental trial was structured as follows. First, the participants saw a fixation cross at the center of the screen for 500 ms. Following a 300 ms blank screen, a distractor sentence was presented at the center of the screen. Participants spent as much time as they needed to memorize the sentence, and pressed the space key when they felt ready. 1000 ms after the key press, another fixation cross appeared on the screen for 300 ms. Following a 200 ms black screen, a picture stimulus appeared on the screen for 5000 ms on 50% of the trials. In that case, participants responded by speaking the target sentence that described the picture stimulus. In the other 50% of the trials, the word REPEAT appeared at the center of the screen, in which case participants responded by repeating back the memorized sentence.

Before the primary trials, participants first studied a booklet containing the picture stimuli that were used in the following experimental session. This booklet also contained the target verb corresponding to each picture. Just like in other PWI studies (e.g., Schriefers et al., 1990), participants studied the (physical) booklet until they felt comfortable with each picture and word. Participants were allowed to spend as much time as they need to study the booklet. The relationship between pictures and words is not arbitrary and is based on previous norms (Szekely et al., 2004). The electronic version of the booklet they saw is available at https://shotam.github.io/CategorySPI/Pictures.pdf.

After this familiarization session, the structure of each trial (illustrated in Fig. 1) was explained to participants. They were instructed to repeat back the sentences or describe the pictures as quickly and accurately as possible. Following the instructions, they performed two practice trials that had the same structure as the experimental trials, using pictures that were not used in the critical trials but were included in the booklet. The experimental session followed this practice session.

For the critical picture trials, the onset latency of the target verb relative to the picture onset was measured using a forced alignment algorithm (Penn Forced Aligner; Yuan & Liberman, 2008) with trial-by-trial hand-correction using Praat (Boersma & Weenink, 2018). All non-target utterances and utterances containing overt fillers were excluded from subsequent analyses (4.6% of all trials). Any trials with a speech onset time of < 300 ms or > 5000 ms (0.5% of the non-erroneous experimental trials), or any trials with onset times more than three standard deviations away from each participant's mean were removed from the data analysis (1.6% of the remaining experimental trials). The production latencies of verbs (the sum of the speech onset latency and production time of preverbal words) were then log-transformed and submitted to statistical analysis. We first measured verb production latency using a text-to-speech alignment algorithm.

All the statistical analyses were conducted using R (R Core Team, 2019), and the lme4 package (Bates, Mächler, Bolker, & Walker, 2015). For the mixed effects model analysis, all the categorical experimental factors were centered (i.e., 0.5, −0.5). The random effects structure was initially maximal in the sense of Barr, Levy, Scheepers, and Tily (2013), but random slopes that caused convergence failures were removed. When simplifying the random effects structure, we removed the random slope term that accounted for the least amount of variance. The structure of the final models is described in the caption of the relevant tables. Aside from the experimental factor, we included the centered trial order as well as its interaction with the experimental factors as a predictor, because we suspected that speakers would be faster to respond in later trials than in earlier trials, and that this trial effect might interact with the relatedness manipulation. We used maximum likelihood ratio tests comparing the model with and without the relevant

Table 1
By-subject means of verb production latency (in ms) by condition (within-subject standard errors) in Experiment 1.

Distractor type	Relatedness	Mean verb latency (se)	Mean onset latency (se)
Noun	Related	1250 (5)	856 (5)
Noun	Unrelated	1250 (7)	849 (6)
Verb	Related	1281 (6)	881 (5)
Verb	Unrelated	1243 (6)	845 (5)

variable, to determine whether trial order or the interaction term should be included.

2.2. Results

In the recall trials (filler trials), participants were able to recall the distractor sentences with 82.2% accuracy. Note that this number reflects speakers' accuracy in the strictest sense. That is, speakers needed to recall the sentences in their exact form to score 'correct.' This shows that participants engaged in the memory task and established a memory encoding of the distractor sentences, as intended, though they may not have been using their verbatim memory in all cases, based on previous evidence suggesting that speakers regenerate sentences from more abstract memory encodings, even in short-term recall tasks (Potter & Lombardi, 1990; Sachs, 1967).

As can be seen in Table 1 speakers took 38 ms longer to start producing the target verb when it was paired with related compared to unrelated verb distractors. However, verb production latency was similar when it was paired with related and unrelated noun distractors (a 1 ms difference). As can be seen in Table 2, this contrast was supported by an interaction between relatedness and distractor category ($p = .01$) as well as by planned comparisons showing a significant effect of relatedness for verb distractors ($\beta = -0.03$, $SE = 0.01$, $|t| = 2.91$, $p < .01$) but not for noun distractors ($\beta = -0.00$, $SE = 0.00$, $|t| = -0.34$, $p = .73$).

The same pattern of results was observed when using speech onset latency as the dependent variable. As can be seen in Table 1 speakers took around 36 ms longer to start speaking when the target sentence was paired with related compared to unrelated verb distractors. However, speech onset latency was similar when the sentence was paired with related and unrelated noun distractors (a 6 ms difference). As can be seen in Table 3, this contrast was supported by an interaction between relatedness and distractor category ($p = .02$) as well as by planned comparisons showing a significant effect of relatedness for verb distractors ($\beta = -0.04$, $SE = 0.01$, $|t| = 3.06$, $p < .01$) but not for noun distractors ($\beta = -0.01$, $SE = 0.01$, $|t| = 0.97$, $p = .34$).

Table 2
Linear mixed effects model estimates of logged verb onset latencies in Experiment 1. The random effects structure included by-subject and by-item random intercepts, and by-item random slopes for relatedness. Removing Trial Order from the model did not change the significance pattern of the other factors.

Term	β	SE	|t|	*p*
Intercept	7.11	0.03	272.55	< 0.00***
Relatedness	−0.02	0.01	2.06	0.051
Distractor category	0.01	0.01	1.24	0.22
Rel. × Dist. Cat	−0.03	0.01	2.44	0.01*
Trial order	−0.00	0.00	24.40	< 0.00***

* means p value is less than .05.
** means p value is less than .01.
*** means p value is less than .001.

Table 3
Linear mixed effects model estimates of logged speech onset latencies in Experiment 1. The random effects structure included by-subject and by-item random intercepts, and by-item random slope for relatedness. Removing Trial Order from the model did not change the significance pattern of the other factors.

Term	β	SE	|t|	*p*
Intercept	6.72	0.03	209.24	< 0.00***
Relatedness	−0.02	0.01	2.46	0.02*
Distractor category	0.01	0.01	1.26	0.21
Rel. × Dist. Cat	−0.03	0.01	2.08	0.04*
Trial order	−0.00	0.00	22.30	< 0.00***

* means p value is less than .05.
** means p value is less than .01.
*** means p value is less than .001.

2.3. Discussion

These results show that interference from conceptually similar words in distractor sentences slowed production of verbs selectively. Slowing occurred when the distractor words were also verbs, but not when they were nouns. This pattern supports the existence of a mechanism that limits competition among across-category words. Furthermore, the current results are more consistent with models that do not assume retrieval of an intermediate category-neutral representation when accessing, for example, *singing* and *whistling*. This is because models that assume the retrieval of an intermediate representation should predict an interference effect between *singing* and *whistling* regardless of their syntactic category, contrary to the current results.

Some critical differences between the SPI task and the PWI task merit discussion. To begin, the relative timing at which the distractor is presented to the participants is different between the current SPI experiment and typical PWI experiments. In the PWI literature, it has been shown that the relative timing of distractor presentation is a critical factor that modulates semantic interference effects (e.g., Schriefers et al., 1990). Indeed, distractors that are presented too early relative to the target stimulus have been shown to cause facilitation rather than interference effects in word-translation interference tasks (Bloem, van den Boogaard, & La Heij, 2004). The word-translation interference task is similar to PWI, except that speakers are asked to translate a written word in a language to another language instead of naming the picture. Speakers in this task, much like in a PWI task, ignore the distractor words.

In the current SPI task, distractor presentation occurs well before picture presentation, but it resulted in an interference effect unlike in Bloem et al. (2004). Thus, we need to explain why the previous PWI-like word-translation experiment showed a facilitation effect while the current experiment showed an interference effect. A potential reason is that, unlike in typical PWI experiments or in the translation interference experiments by Bloem et al. (2004), we asked speakers to memorize the distractors. As we hypothesized in the Introduction, this may prevent the activation of the lexical representation of the distractor from decaying over time, thereby making the distractor word a long-lasting source of competition for target production. This is consistent with the explanation by Bloem et al. (2004) that conceptually related distractors that are presented too early cause facilitation because lexical activation decays more quickly than conceptual activation, which causes priming in subsequent target production (i.e., facilitation). Under this view, the difference arises due to the *task relevance* of the distractor. In Bloem et al. (2004), distractors were irrelevant to the task at hand so there was no reason for speakers to keep them active. In contrast, the distractors in the current experiment were task-relevant because speakers needed to say them 50% of the time. Thus, speakers had reasons to keep distractors active until they saw target pictures, only after which they actively suppressed the distractor activation. Hence, despite the timing difference, we argue that the SPI task is consistent with the model of distractor effects built to explain the results of PWI experiments.

In addition, unlike a typical PWI task, the current experiment involved sentence-level production, in which the target word appeared last in the target sentence. This difference might raise an additional concern that the timing of lemma selection may be too late for the distractor to cause interference. However, there are two reasons to believe that this concern is unwarranted. First, the results show that speakers showed a verb semantic interference effect in the sentence onset latency measure, suggesting that speakers select verbs early before sentence onsets. In addition, we found independent evidence that speakers planned the verbs early in target sentence production. A post-hoc mixed effects analysis, in which the lexical frequency of the verb (as well as all other experimental factors) was included as a fixed effect, suggests that the lexical frequency of verbs was a significant predictor of speech onset latency ($\beta = -0.04$, $SE = 0.01$, $t = -3.01$, $p < .01$). This result also suggests that speakers planned the verb early, before utterance onset. Thus, we have converging evidence that speakers select target verb lemmas early before the sentence onset, on average within 850 milliseconds after the onset of picture presentation. Therefore, the potential concern that the timing of verb selection was too late for the distractor word to cause interference in the current experiment is empirically unsupported.

The evidence for early verb planning in the current experiment might seem in conflict with previous studies that suggest that verbs do not need to be planned before the sentence onset (Schriefers, Teruel, & Meinshausen, 1998), at least not when the sentence-initial constituent is an agentive subject (Momma et al., 2018; Momma, Slevc, & Phillips, 2016). However, there are many relevant differences between the current experiment and previous experiments. First, in the current experiment, verbs were the first content word that speakers needed to select. Second, in the current experiment, the subject noun phrase was a pronoun, which is very short phonologically. Speakers may have planned the later-coming words in advance when the initial word was phonologically short (Griffin, 2001). Third, in the current experiment the subject noun and the verb constituted a single prosodic unit, which may be a unit of planning (Schnur, Costa, & Caramazza, 2006). All these differences are in a direction that encourages speakers to plan sentence-final verbs early, before speech onset. Thus, we argue that there is no clear conflict between the current finding and previous studies.

In Experiment 1, we did not observe any facilitatory effects (although there was a numerical facilitation due to the distractor type in the unrelated noun distractor conditions, compared to the unrelated verb distractor conditions), and the lack of any facilitatory effect may seem surprising. For example, one might expect that the verb distractor sentences (e.g., *Mary told the doctor that she is persistently coughing*) should facilitate production of the target verb, as compared to noun distractor sentences (e.g., *Mary told the doctor about her persistent coughing*). This is because of past evidence of grammatical class priming Melinger and Koenig (2007) (*coughing* as a verb primes *whistling* as a verb), syntactic priming (Bock, 1986), and semantic facilitation effects from cross-category distractors (Mahon, Costa, Peterson, Vargas, & Caramazza, 2007, Bloem et al., 2004, but see Vigliocco et al., 2004). However, considering the design and materials of the current experiment, the lack of these possible facilitatory effects is not particularly surprising.

First, we should note that these potential facilitatory effects from grammatical class repetition or structural repetition were confounded with other factors in the current study, such as the length and complexity of distractors. Thus, the current result is not informative about whether speakers experience a grammatical class priming effect or syntactic priming effect.

Second, it is not clear if we should expect any of the facilitative effects mentioned above in this experiment. Every distractor sentence contained nouns and verbs multiple times, so any grammatical class priming effect should actually be present in all conditions. Certainly, the size of the hypothetical grammatical class priming effect might be different if one assumes that grammatical class priming is cumulative, because verb distractor sentences contain one more verb than noun distractor sentences. But it would not be easy to detect such a subtle difference. Moreover, if the grammatical class priming effect is not cumulative, then we should not expect any difference between verb and noun distractor conditions in terms of grammatical class priming. In addition, in the current experiment target production did not involve a syntactic choice, unlike in other syntactic priming studies, so we should not necessarily expect a syntactic priming effect.

Finally, we did not observe any facilitative effect from related distractors in the noun distractor conditions. Mahon et al. (2007, Experiments 5ab, 6, 7ab) found that distractors that are conceptually related to the target (e.g., *futon* for a target *bed*) facilitated production relative to less related distractors (e.g., *chair*). Also, as discussed above, Bloem et al. (2004) found in a PWI-like word translation task that, when distractors were presented sufficiently early (400 ms before the target stimulus), production was facilitated by conceptually related distractors (e.g., *goat* for the target *pig*) relative to unrelated distractors (e.g., *glass*). These facilitatory effects in PWI and PWI-like translation tasks are often hypothesized to arise from spreading activation at the conceptual level of representation from distractor to target concepts (e.g., Abdel Rahman & Melinger, 2009). One may argue that, if syntactic category mismatch weakens or eliminates the interference effect, we should expect to see the facilitatory effect due to conceptual priming, contrary to the current results.

There are at least three possible reasons that we did not observe a facilitatory effect in the current experiment. First, it is possible that the facilitatory effect is simply not robust. Mahon et al. (2007) found a reliable facilitatory effect in their Experiments 5a and 5b, but their Experiments 6 and 7a did not show a significant effect in the items analysis, and the facilitatory effect in Experiment 7a was only found in the subjects analysis in one of the three SOA conditions (only in the −160 ms condition). They also found no facilitatory effect in their Experiment 7b, where 0 ms SOA was used. Furthermore, Vigliocco et al. (2004) found precisely the opposite results; they found that, as the semantic similarity between a categorically related distractor and target words increased (e.g., for a target word *bake*, *grill* was the semantically closest, *cook* the second closest, *eat* the third closest, and *drop* the furthest) speech onset latency also increased (i.e., they observed graded interference effects). Thus, there is considerable variability in the robustness of the facilitatory effect, and it is not clear if we should expect to see a facilitatory effect in the current experiment.

Second, it is possible that the facilitatory effects found in some previous experiments reflect strategic expectations, not automatic spreading activation. In word recognition research, there are two potential accounts for facilitatory effects in lexical processing: automatic spreading activation and strategic expectations. In fact, expectation accounts may account for a large portion of the facilitatory effect in the lexical decision task (e.g., Lau, Holcomb, & Kuperberg, 2013). According to this view, the facilitatory effect of a conceptually related distractor in PWI-like tasks may not be primarily due to automatic spreading activation, but is instead to a large extent due to speakers' strategic expectations (e.g., speakers expect to see *dog* when *cat* is presented). This possibility is naturally compatible with the pattern that the facilitatory effect is only observed when the distractor is presented sufficiently early, because generating expectations takes time. Bloem et al. (2004) found a facilitatory effect when the distractors were presented 400 ms before the target stimulus, but Bloem et al. (2004) did not find such a facilitatory effect when the distractors were presented 200 ms before the target stimuli. Though the timing was different, Mahon et al. (Experiments 7ab) found a facilitatory effect only when the distractors were presented 160 ms before the target pictures, and not when they were presented at the same time as the target pictures or 160 ms after the target pictures. In the current SPI task, strategically expecting a particular target based on a distractor is likely much harder than in the PWI task, because the relevant distractor was embedded in a complex sentence context. This may explain why some previous studies in PWI-like tasks showed a facilitatory effect from related distractors while we failed to see it in the current experiment.

Third, it is possible that there was indeed a weak facilitatory effect, but we failed to see it because of (a) a lack of sufficient power or (b) a weak inhibitory effect canceled out by a weak facilitatory effect in the across-category conditions. Those possibilities cannot be ruled out by the current results. However, any of the above described possibilities do not undermine the main implication of the current study that syntactic category has a competition-limiting function.

In Experiment 1, there were some unavoidable confounds between the verb and noun distractor conditions. For example, the distractor sentences in the noun distractor conditions were systematically shorter, and less complex (in terms of number of clauses). These confounds were unavoidable by design; a verb distractor word by definition creates an additional clause, so it is impossible to match the number of clauses while also matching the other parts of the distractor sentences as much as possible. It is possible that this difference somehow interacted with the semantic interference effects. Thus, Experiment 2 tested whether the reverse pattern of interference effects can be obtained when speakers produce nouns instead of verbs.

3. Experiment 2

Experiment 2 aimed to test whether the production of nouns can be interfered with by noun distractors but not by verb distractors, using a maximally similar design to Experiment 1. However, this is not as easy as it first seems. There are few natural situations in which speakers would say a noun version of *whistling*. In order to elicit the noun version of *whistling* from speakers, we placed a colored square in the corner of each target picture (see Fig. 4 for an example), and changed the instructions of the picture description component of the task to the following:

*Imagine yourself in a hypothetical world where you perceive a color for each action. You know that some people, specifically people who have what's called synesthesia, perceive colors for things like numbers and letters. In your case, you perceive a color for each action. Depending on the kind of action and depending on who does it, you perceive different colors (specifically, the color you see in the right lower corner of the pictures). Your task is to report the color of each action, using a full sentence of the form X's Ving is red/blue/*etc.

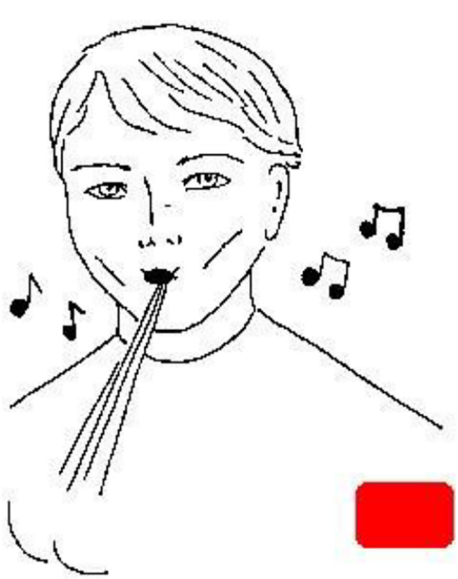

Fig. 4. An example picture stimulus in Experiment 2.

After a couple of practice trials, these instructions reliably elicited sentences containing gerundive nominals, e.g., *her whistling is red*. Using this method, Experiment 2 tested whether noun distractors selectively interfered with noun production.

In Experiment 2 we measured noun production latency rather than speech onset latency, for the same reason we measured verb production latency in Experiment 1.

3.1. Method

3.1.1. Participants

Twenty-four University of Maryland undergraduate students participated in Experiment 2 for course credit. Informed consent was obtained from all participants prior to the experiment. None had participated in Experiment 1.

3.1.2. Materials

An example picture stimulus is presented in Fig. 4. The pictures (except the colored square) and distractor sentences were identical to those used in Experiment 1.

3.1.3. Procedure and analysis

The experimental procedures and analyses were identical to Experiment 1, except that participants were given different instructions for the picture description component of the task, as described above. Non-target utterances and utterances containing overt fillers were excluded, just like in Experiment 1, and accounted for 10.4% of the experimental trials.

3.2. Results

In the recall trials (filler trials), participants were able to recall the distractor sentences with 81.0% accuracy. Again, note that this number reflects speakers' accuracy in the most strict sense. That is, speakers needed to recall the sentences in their exact form to score 'correct.'

As can be seen in Table 4, speakers took 41 ms longer to start producing the target noun when paired with related compared to unrelated noun distractors. However, noun production latency was similar when paired with related and unrelated verb distractors (a − 12 ms difference). This was supported by an interaction between relatedness and distractor category (Table 5) as well as by planned comparisons showing a significant effect of relatedness for noun distractors ($\beta = -0.04$, $SE = 0.01$, $|t| = 2.34$, $p = .03$) but not for verb distractors ($\beta = 0.00$, $SE = 0.01$, $|t| = 0.29$, $p = .78$). As in Experiment 1 latencies also decreased over the course of the experiment, as reflected in a main effect of trial order (Table 6). No interaction involving trial order was found, so those interactions were removed from the final model. As predicted, the pattern of interaction effect was exactly the opposite of the pattern of Experiment 1, suggesting that when the target word is a noun only noun distractors are effective in inducing the semantic interference effect. In addition, unlike in Experiment 1, there was a main effect of Distractor Category, but this effect is hard to interpret due to the presence of a significant interaction.

Table 4
By-subject means of subject noun production latency and sentence onset latency (in ms) by condition (within-subject standard errors) in Experiment 2.

Distractor type	Relatedness	Mean verb latency (se)	Mean onset latency (se)
Noun	Related	1192 (7)	995 (7)
Noun	Unrelated	1151 (6)	960 (7)
Verb	Related	1141 (9)	952 (9)
Verb	Unrelated	1153 (8)	958 (8)

Table 5
Linear mixed effects model estimates of logged verb onset latencies in Experiment 2. The random effects structure included by-subject and by-item random intercepts, and by-item random slope for relatedness. Removing Trial Order from the model did not change the significance pattern of the other factors.

Term	β	SE	|t|	p
Intercept	7.03	0.03	219.43	< 0.01***
Relatedness	−0.02	0.01	1.41	0.17
Distractor category	−0.02	0.01	2.64	0.008**
Rel. × Dist. Cat	0.04	0.02	2.34	0.02*
Trial order	−0.00	0.00	13.62	< 0.00***

* means p value is less than .05.
** means p value is less than .01.
*** means p value is less than .001.

Table 6
Linear mixed effects model estimates of logged speech onset latencies in Experiment 2. The random effects structure included the by-subject and by-item random intercepts. Removing Trial Order from the model did not change the significance pattern of the other factors.

Term	β	SE	|t|	p
Intercept	6.85	0.03	195.65	< 0.00***
Relatedness	−0.02	0.01	2.09	0.04*
Distractor category	−0.02	0.01	2.53	0.01*
Rel. × Dist. Cat	0.04	0.02	2.00	0.045*
Trial order	−0.00	0.00	12.04	< 0.00***

* means p value is less than .05.
** means p value is less than .01.
*** means p value is less than .001.

The same pattern of results was observed when using speech onset latency as the dependent variable. As can be seen in Table 4, speakers took 35 ms longer to start the sentence when it was paired with related compared to unrelated noun distractors. However, sentence onset latency was similar when paired with related and unrelated verb distractors (a 7 ms difference in the opposite direction). As can be seen in Table 6, this contrast was supported by an interaction between relatedness and distractor category as well as by planned comparisons showing a significant effect of relatedness for noun distractors ($\beta = -0.04$, $SE = 0.02$, $|t| = 2.36$, $p = .03$) but not for verb distractors ($\beta = -0.00$, $SE = 0.01$, $|t| = -0.02$, $p = .99$).

3.3. Discussion

In Experiment 2, interference from conceptually similar words in distractor sentences slowed the production of nouns only when the distractor word was also a noun. This provides additional evidence for a mechanism that limits interference between words in different syntactic categories. Also, for the reasons discussed in the Introduction, the current results together with the results of Experiment 1 are more consistent with models that do not assume retrieval of intermediate representations in accessing complex gerundive nominals. If accessing *whistling* as a noun involves an intermediate step of accessing the category-less root √*whistle*, or the verb *whistle*, then some degree of interference is expected regardless of the eventual category of the critical words in the distractor sentences. This was not the case in Experiment 2 or in Experiment 1. Thus, these results show the effectiveness of the syntactic category constraint on morphologically complex nominals that are internally verbs, and thus they support models of lexical access that do not assume that initial category-less representations or some category-changing derivational processes are involved in word retrieval processes.

The combination of verb-specific and noun-specific interference found in Experiments 1 and 2 further shows that any potential confounding differences in naturalness, length, and complexity between distractor conditions in Experiment 1 were unlikely to be the cause of the selective interference pattern. This is because Experiment 2 used the exact same set of distractors (and pairing with target pictures), and thus had exactly the same potential confounds, yet showed the opposite pattern of results (see Fig. 5 below).

We found a main effect of Distractor category in Experiment 2 but not in Experiment 1, but this difference is unlikely to be meaningful as Experiment 1 also showed the same numerical trend. Also, a main effect of Distractor category can be caused by factors like the length of distractor sentence, and therefore it is not of theoretical interest here.

4. General discussion

The current experiments yielded two main results. Experiment 1 demonstrated that the semantic interference effect in verb production is

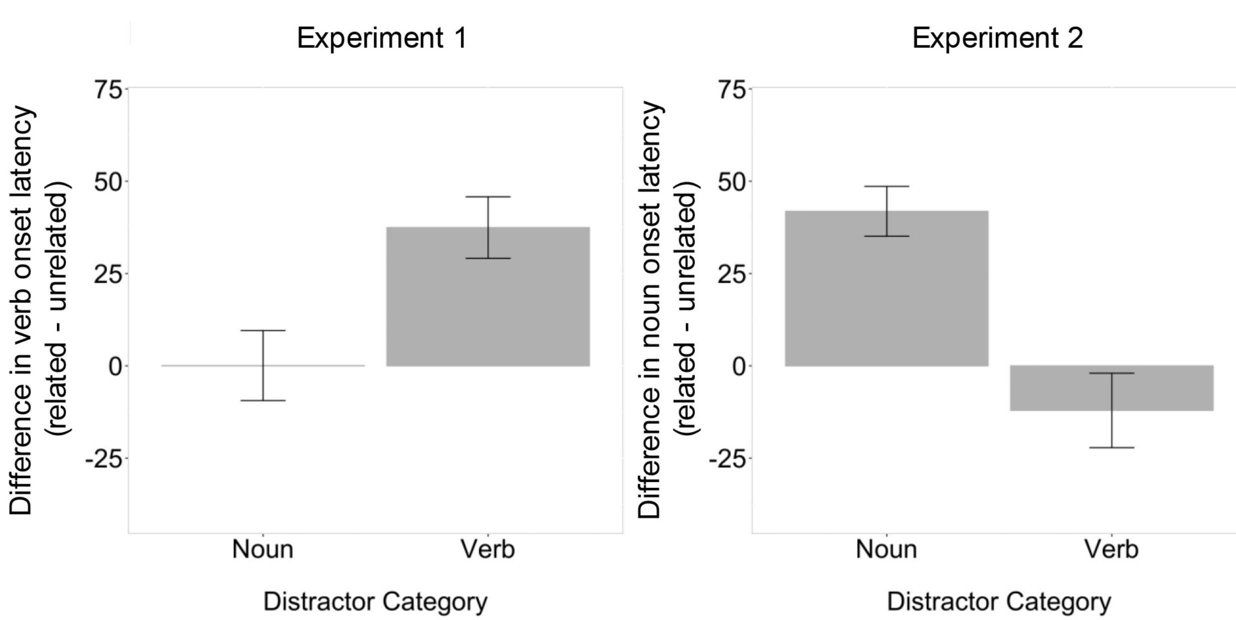

Fig. 5. Interference effects in Experiment 1 and Experiment 2.

observable only when the distractors are also verbs. Experiment 2 showed that the semantic interference effect in noun production is observable only when the distractors are also nouns. The pattern of interference in Experiments 1 and 2 is summarized in Fig. 5. This interference pattern suggests that competition in lexical access is restricted to words from the same syntactic category.

4.1. Theoretical implications

The current results have five theoretical implications above and beyond previous studies on speech error patterns (Garrett, 1975; Nooteboom, 1980) and computational modeling (Dell et al., 2008) that explicitly examines the interaction between syntactic category information and conceptual information in lexical selection in sentence-level production. In the following discussion we assume the model of word production by Levelt et al. (1999), in which lemma selection processes are assumed to precede morphophonological encoding processes. Because our main effects of interest (i.e., the semantic interference effect) is normally assumed to be an effect on lemma selection (e.g., Schriefers et al., 1990), we limit our discussion to how speakers select appropriate lemmas. Thus, all of our claims below should be construed as strictly about lemma access, not about word-form encoding processes.

4.1.1. The interaction between conceptual and syntactic category information

The current results suggest that speakers can use syntactic category information to constrain the effect of conceptual similarity in the lemma selection process. This result constrains previous models of production that explicitly discuss how syntactic category information interacts with conceptual information to enable lexical selection, such as Dell's syntactic traffic cop model (Dell et al., 2008).

In Dell's syntactic traffic cop mechanism, the syntactic category constraint is implemented as nodes representing syntactic categories separately from individual words (noun, verb, adjective) and connecting to each of the lexical nodes (*cat*, *dog*, *bark*, etc.). The connection weights between the syntactic category nodes and lexical nodes can be positive or negative. The state of the network changes depending on processing context such that, for example, the noun category node is activated when a determiner was the last word that the model processed. The activation of category nodes enhances all the lexical nodes connected with positive weights, and it inhibits all the lexical nodes connected with negative weights. For example, after a determiner, the noun category node is activated, which in turn activates lexical nodes that are connected to the noun category node with positive weights (e.g., *dog*). In contrast, lexical nodes that are connected with negative weights (e.g., *bark*) are inhibited. This reduces the effect of competition from inhibited lexical nodes, that is, words in different syntactic categories than the target.

The current results constrain the parameter space of this syntactic traffic cop model. One possible configuration of this model is such that the activation from syntactic category representation and conceptual representation add up. This can explain, as Dell et al. suggested, the pattern that Vigliocco et al. (2005) observed in their PWI experiments discussed in the Introduction, namely the additive effect of conceptual similarity and syntactic category similarity between the distractor and target words. However, the current results are not easily compatible with this particular set up of the model, because we found that the effect of conceptual similarity was only observed in the presence of syntactic category overlap. Thus, one of the theoretical contributions of the current study is that it constrains how syntactic and lexical information interacts in explicit models like Dell et al. (2008), or more broadly any models that aim to capture the relationship between syntactic category information and lexical access.

4.1.2. The nature of lexical memory

The current results also suggest that speakers are capable of storing and accessing even complex words, together with their syntactic category information. Assume the contrary view that morphologically complex words are always computed *on the fly* because they are not stored in lexical memory. This view can be called a *computation-only view*. For example, when speakers try to say *running* as a noun (as in Experiment 2), speakers need to (a) select a verb stem (*run* [v]), (b) select a function morpheme that nominalizes it (e.g., -ing) and (c) combine the stem (*run* as a verb) and nominalizing *-ing* to create *running* as a noun. Alternatively, under the view that the stem morphemes are categorically underspecified (Barner & Bale, 2002; Garrett, 1975; Pfau, 2009), speakers need to (a) select a category-less stem (*run*), (b) select a verbalizer morpheme, and (c) combine the two representations to create the verbal version (*run* [v]), (d) select the nominalizing *-ing*, and finally (e) combine *run* as a verb with the nominalizer *-ing*. Either way, the computation-only view presupposes that *run* as a verb, or category-less *run*, is initially selected when producing *running* as a noun. Therefore, under the computation-only view, distractor verbs (like *walking* as a verb) should cause the same amount of semantic interference than distractor nouns (like *walking* as a noun), even when speakers say *running* as a noun. Contrary to this prediction, the results of Experiment 2 showed that noun distractors are more effective at inducing semantic inference for nouns as targets. Thus, speakers in the current experiment skipped the process of selecting *run* as a verb (or a category-less version of it) when speaking *running* as a noun. A plausible way to do so is to directly access a morphologically complex representation of nominal gerunds, stored in memory as a noun.

Admittedly, this argument involves assumptions that may turn out to be false. Thus, this argument about morphological representation should be evaluated cautiously, together with past and future work on morphological representation that deals with the problem of (non-) decomposition more directly. Also, it is important to emphasize that the current results do not suggest that every single complex word is always stored in lexical memory and accessed via a single retrieval process. Indeed, the view that every single complex word is always stored in memory would be highly unlikely, because speakers also need to be able to retrieve smaller units (i.e., morphemes) and derive complex words from them, rather than only retrieve complex words as a whole. Otherwise, morphological productivity (e.g., speakers' ability to use a noun as a novel verb) would remain unexplained. Thus, we argue that speakers can (a) actively project the syntactic category of an upcoming word, and (b) use it to (selectively) access a memory entry whose syntactic category matches the projected category.

The current question is closely related to the issue of morphological (de-) composition in comprehension and production. There is a rich history of research in comprehension (Caramazza, Laudanna, & Romani, 1988; Laudanna, Badecker, & Caramazza, 1992; Lukatela, Gligorijević, Kostić, & Turvey, 1980; Marslen-Wilson, Tyler, Waksler, & Older, 1994) and in production (Anshen & Aronoff, 1988; Garrett, 1975; Zwitserlood, Bölte, & Dohmes, 2000, 2002) investigating whether comprehenders and speakers (de-)compose morphologically complex words. As discussed above, the results of Experiment 2 provide evidence for the view that morphologically complex words can be stored and accessed as a unit. However, it is important to note that this evidence is compatible with the view that speakers do sometimes compose morphologically complex words in production. The contrary view that speakers do not compose morphologically complex words at all is implausible, given the fact that morphology can be completely productive. It is also important to note that complex words involving derivational and inflectional morphemes may behave differently, because, for example, inflectional morphemes are generally highly productive and yield predictable interpretations, whereas derivational morphemes are often less productive and can have idiosyncratic meanings. Thus, even if the current results provide some evidence for the non-decomposition of words involving derivational morphology, they do not necessarily generalize to words involving inflectional morphology (Anshen & Aronoff, 1988; Laudanna et al., 1992).

4.1.3. The status of syntactic category

The current results also suggest that speakers can represent abstract syntactic category largely independently of conceptual contents. This is in line with the past behavioral studies (Melinger & Koenig, 2007), speech error studies (Dell et al., 2008; Garrett, 1975; Nooteboom, 1980), neuroimaging studies (Bedny, Caramazza, Grossman, Pascual-Leone, & Saxe, 2008), and neuropsychological studies (Caramazza & Hillis, 1991). But there remains some skepticism about genuine noun-verb distinctions at the neural level that is dissociable from semantic classes (e.g., Moseley & Pulvermüller, 2014). In the current experiment, unlike in the speech error studies which tend to be observational, the semantic content of nouns and verbs were matched to the greatest extent possible. We acknowledge the possibility that there may be subtle semantic differences between the meaning of progressive and gerundive *-ing* forms. However, we do not regard this as a serious alternative semantic account of our results. The semantic contribution of the lexical root is identical in both cases, and there is no serious suggestion that subtle semantic differences between *singing* (noun) and *singing* (verb) results in measurable differences in how much semantic similarity each word has with *whistling* (noun) and *whistling* (verb). Thus, the current results can be used as evidence that syntactic category that is dissociable from semantic content is represented and used in production. Of course, this does not necessarily mean that syntactic category information is always represented whenever speakers utter a word. Our task was a sentence-level production task, and that might have magnified the effect of syntactic category, consistent with previous claims (Vigliocco, Vinson, Druks, Barber, & Cappa, 2011). Consequently, the role of syntactic category in single-word production might be smaller than the effects reported here.

4.1.4. Top-down structure building

In order for syntactic category to constrain lexical access, speakers must be able to represent syntactic category before the point in the lexical selection process where competition occurs. This view contrasts with some versions of well-known lexicalist models of production (e.g., Bock & Levelt, 1994), in which syntactic categories are projected from words in a *bottom-up* fashion. Under the account proposed here, speakers are capable of projecting syntactic categories using their knowledge of phrase structure, in a *top-down* fashion, before knowing what words fill those projected structural positions. Of course, this does not mean that speakers always project categories ahead of lexical access, or that the syntactic structure of an entire clause needs to be planned in advance (e.g., F. Ferreira, 2000; Garrett & Newmeyer, 1988). In our view, a single sentence production process can probably involve interleaving procedures of both bottom-up and top-down structure building (Momma & Phillips, 2018). Also, there may be many situations where speakers are more certain about the content they want to express than about the structure of their intended expression. However, speakers can use category information to constrain lexical access. This *top-down* method of building syntactic structure may help speakers to minimize interference.

The involvement of top-down structure-building may at first seem necessary for explaining the syntactic category constraint in speech errors (Dell et al., 2008; Garrett, 1975; Nooteboom, 1980). However, it is possible to capture the syntactic category constraint without assuming the involvement of top-down syntactic structure building. For example, in line with the view that the availability of a word determines the appropriate structure rather than the other way around (e.g., V. S. Ferreira, 1996), the syntactic category constraint may be a constraint on integration processes between structural and lexical representations. Under this view, lexical selection is relatively unconstrained by structural representations, to allow more flexible, and hence more fluent production (e.g., Bock, 1982; V. S. Ferreira, 1996; Wardlow-Lane & Ferreira, 2010). Speakers still respect the syntactic category constraint because verbs, even when selected, cannot be grammatically integrated with structures with a determiner as their previous word and successful integration is a prerequisite for articulation. However, speakers can still perform integration processes relatively flexibly, for example using derivational morphological processes. This approach, however, predicts that words that do not share the syntactic categories should still compete with each other, contrary to the current results.

Thus, under the assumption that the semantic interference effect we observed in the current study reflects increased competition in lexical selection (but see below for an alternative interpretation), the current result is more compatible with the view that syntactic category information is projected before and independently of lexical selection. The current results show that a syntactic category mismatch can block the semantic interference effect. In this sense, the current results complement the speech error evidence and help us characterize the time-course of structural and lexical processes in sentence-level production. Note again, however, that speakers might not use a top-down structure building method in all task environments. Speakers may build structures in a bottom-up fashion when they are uncertain about sentence structure prior to lexical access. Nevertheless, the current experiments show that speakers are capable of building syntactic structures top-down.

4.1.5. Advance planning and syntactic category

Finally, the current claim is relevant to how far in advance speakers plan lexical information in sentence production (Allum & Wheeldon, 2009, 2007; Konopka, 2012; Meyer, 1996; Schriefers et al., 1998; Smith & Wheeldon, 1999; Wagner, Jescheniak, & Schriefers, 2010). It is generally agreed that the extent of advance lexical planning is limited, primarily because it is costly to hold multiple words in working memory simultaneously. It has been assumed that the cost of advance lexical planning originates from similarity-based interference (De Smedt, 1996), and similarity-based interference is thought to be the primary reason for speakers to try to maximize synchronization between planning and articulation of each word by avoiding advance lexical planning (De Smedt, 1996; Iwasaki, 2010). Given the current view that only words from the same syntactic category compete with each other, however, such a view may be too simplistic. For example, speakers may be able to plan a verb before uttering a noun to ensure that the grammatical status of preverbal nouns conforms to the selectional requirements of the verb, without increasing memory cost (Momma et al., 2016, 2018). This may, therefore, help to explain why speaking can be both relatively fluent and grammatically robust.

4.2. Relationship to speech errors

The current results converge well with typical patterns of whole-word exchange errors. As discussed in the Introduction, whole-word exchange errors obey the well-known syntactic category constraint, e.g., *The frisbee caught the dog* is a likely error, but *The caught dog the frisbee* is a highly unlikely error (for the intended: *The dog caught the frisbee*). This pattern is exactly as expected if lexical competition does not occur between words belonging to different syntactic categories. Thus, the current results provide additional evidence for the view that the syntactic category of whole words is used in lexical access to limit competition, even in non-erroneous speech, and thus they show that the same mechanism can explain latency data in experimental settings (as shown here) and speech error data in naturalistic settings. The current data also allow us to understand the processing locus of the syntactic category constraint. As discussed above, the syntactic category constraint could, in principle, arise in lexical selection and/or lexical integration. Under the assumption that the semantic interference effect is an effect on lexical selection, the current results favor the view that the syntactic category constraint arises from the nature of the lexical selection mechanism.

One interesting type of errors that are seemingly in conflict with the current claim is stranding errors. Stranding errors are a type of exchange error in which morphemes (often root morphemes) exchange, leaving the inflectional and derivational morphology behind. For example, speakers might erroneously say *trucked the record* instead of *recorded the truck*. An interesting property of this type of error is that it does not obey the syntactic category constraint (see Pfau, 2009, but see also V. S. Ferreira & Humphreys, 2001). One explanation of the contrast between whole word exchange and stranding errors comes from Garrett (1975), who considered stranding errors to arise at a "positional" level of processing, in which serial order and phonological specification of words are determined. This positional level of processing is temporally preceded by the "functional" level of processing, in which the structural positions of words are determined, and in which whole-word exchange errors occur. Crucially, we argue that the functional processing stage is the only stage where the syntactic category constraint applies. As Garrett argued, stranding errors occur at the later positional processing stage. As a consequence, stranding errors are less sensitive to the syntactic category constraint than whole-word exchange errors, because words are already bound to the structural position before the positional process begins. Given this view, the existence of errors that violate the category constraint is only seemingly in conflict with the existence of the category constraint. The current results are therefore compatible with standing error patterns under the standard assumption that the type of interference effects seen in the current experiments occurs at the functional level of processing, which includes lemma access (e.g., Bock & Levelt, 1994).

4.3. Alternative interpretations and limitations

So far, we have discussed the current results under the working assumption that interference effects reflect increased competition in lexical access. We regard this as a simple and well-motivated interpretation of the effects, given the previous literature on picture-word interference and the results of our single-word production experiment (see Appendix).

However, it is possible to interpret the interference effects in different ways. One such possibility is that the interference effect in the SPI task reflects a post-lexical integration process, i.e., the process by which the selected word is integrated into the overall representation of a sentence. Under this interpretation, the post-lexical integration process is constrained by syntactic category in such a way that the integration process is easier when competitor words are of a different syntactic category. The semantic interference effect we observed in the current SPI task has not been explored in detail yet, so this alternative interpretation cannot be ruled out. However, it should also be noted that this post-lexical integration account is not parsimonious. First, it does not explain why a similar interference effect was also observed in the single word naming task reported in the Appendix, just like in PWI tasks involving single-word production. Second, it would need to claim that lexical selection itself is immune to the interference from distractors, and that semantic interference effects arise at a later stage. Furthermore, we believe that the current experiments offer useful results even under this alternative interpretation. First, they would show that syntactic category somehow limits post-lexical competition. Second, they would show that this integration process targets whole words, not their parts. These are as theoretically interesting as the claim that we made above, based on the assumption that the interference effect reflects increased lexical competition.

Also, we acknowledge that the current results leave open the question of how broadly syntactic category constrains lexical selection, when we move beyond the specific design used here. The type of sentences that speakers produce in the current experiment is limited and relatively homogeneous in structure. Thus, the current experimental task might have encouraged speakers to use syntactic category information to constrain lexical access because speakers could be confident about the structure of sentences before knowing what specific words to use in the current experiments. Thus, it is possible that speakers might not use syntactic category to constrain lexical access in everyday speech. However, the main point of the current study is to show that humans have a mechanism that, in principle, allows them to use syntactic category to block well-known effects of semantic similarity. The current results suggest that they do, and the naturalistic data from speech errors also point to the same conclusion.

4.4. Methodological contributions

In this article we introduced the Sentence Picture Interference paradigm. To our knowledge this task has not been used previously. A major advantage of this task is that it allows the manipulation of contextually determined properties of distractor words, e.g., whether a word is noun or a verb, whether a word is subject or object, whether a word receives agent or patient role, etc. Thus, this task allows the investigation of how these contextually determined properties of words affect lexical access. This could not easily be done with traditional picture-word interference tasks, for the reasons we discussed in the Introduction. Thus, we hope that this method will prove useful in bridging the existing gap between single word production research and sentence production research.

5. Conclusion

Speakers need to somehow efficiently manage retrieval interference in order to speak relatively fluently and without too many errors. In order for this to happen, speakers need to organize and access their lexical memory in a way that minimizes memory interference caused by lexical competition. In this article, we reported experimental evidence suggesting that syntactic category plays a major role in limiting competition in sentence production. We argued that speakers access complex word representations together with their syntactic category information. We speculated that this might be part of the reason why native speakers are able to speak fluently and grammatically at the same time, at least most of the time.

Author contribution

Shota Momma: Conceptualization, Methodology, software, formal analysis, investigation, resources, data curation, writing, visualization, project administration, investigation **Julia Buffinton**: writing - review & editing, investigation **Colin Phillips**: Conceptualization, supervision, funding acquisition, writing - review & editing **Robert Slevc**: Conceptualization, supervision, writing - review.

Acknowledgements

This research was supported in part by National Science Foundation grant BCS-0848554 to Colin Phillips and National Institutes of Health grant R01-HD051030 to Victor Ferreira.

Appendix A. Testing the SPI task

In this appendix we report an experiment that tests the effectiveness of the sentence-picture interference (SPI) task. To our knowledge, this is a novel task that has not been used before. Though the current task is similar to the widely used picture-word interference task, there are several important differences. For instance, in the SPI task distractor stimuli are not ignored but memorized. Distractor stimuli are also presented well before picture presentation in SPI, unlike in PWI where they appear simultaneously. Given the clear differences between the current SPI and previous PWI tasks, we aimed to test whether the SPI task is sensitive to the semantic interference effect that is usually seen in the PWI task. In order to do so, we used the object pictures (like pictures of *fox*, *dog*, *apple*, *lemon* etc.) that are stereotypically used in PWI studies, unlike in Experiments 1 and 2, in which action pictures were used. To the extent that this study shows a semantic interference effect, we can be confident that the SPI task is effective for measuring the semantic interference effect, the effect of interest in Experiments 1 and 2.

A.1. Methods

A.1.1. Participants

Twenty-four University of Maryland undergraduate students participated for course credit. Informed consent was obtained from all participants prior to the experiment. Two participants were excluded from the analysis due to recording failures, and one participant was excluded because the participant needed to leave before the experiment completed. The data from the remaining twenty-one participants were analyzed.

A.1.2. Materials

Twenty-four object pictures and corresponding target words were selected from the International Picture Naming Database (Szekely et al., 2004).

Twenty-four distractor sentences were constructed and paired with pictures such that the last word of the sentence was semantically related to the target word of the pictures. The degree of semantic relatedness was determined based on intuitive judgment and then verified by assessing cosine distance in Latent Semantic Analysis (LSA; Landauer & Dumais, 1997). The related distractor sentences were re-paired with another target to create unrelated target-distractor pairs. That is, the sets of related and unrelated distractor sentences were identical, with conditions varying only in the relationship between distractor sentences and the following pictures. The mean LSA cosine distance between the target and the last word of the distractor sentence was 0.40 (sd = 0.23) in the related pairs and 0.06 (sd = 0.07) in the unrelated pairs (two-tailed *t*-test; $t(46) = -6.92$, $p < .001$). A related distractor sentence and an example distractor sentence and a target picture for each related and unrelated condition are provided in Fig. 3.

Because the sets of related and unrelated distractor sentences were identical, differences in sentence complexity, plural/singular differences, etc., cannot explain any difference between related and unrelated distractor conditions. The related words in the distractor sentences were also used as target words, which has been suggested to maximize the chance of obtaining a semantic interference effect (Roelofs, 1992).

A.1.3. Procedure and analysis

Each experimental trial was structured as follows. First, the participants saw a fixation cross at the center of the screen for 300 ms. Following a 200 ms blank screen, a distractor sentence was presented at the center of the screen. Participants spent as much time as they needed to memorize the sentence and pressed the space key when they felt ready. Right after the key press, another fixation cross appeared on the screen for 300 ms. Following a 200 ms blank screen, a picture stimulus appeared on the screen for 5000 ms on 50% of trials. In that case, participants responded by saying the target word that corresponded to the picture stimulus. In the other 50% of trials, the word *repeat* appeared at the center of the screen, in which case participants responded by repeating back the memorized sentence.

Before the primary trials, participants first studied a booklet containing the picture stimuli that were used in the following experimental session. This booklet also contained the target word corresponding to each picture. Just like in other PWI studies (e.g., Schriefers et al., 1990), participants studied the booklet until they felt comfortable with each picture and word. The relationship between pictures and words is not arbitrary and is based on previous norms (Szekely et al., 2004). The electronic version of the booklet used is available at https://shotam.github.io/CategorySPI/Pictures.pdf. After this familiarization session, the structure of each trial (illustrated in Fig. 1) was explained to participants. They were instructed to repeat back the sentences or describe the pictures as quickly and accurately as possible. Following the instructions, they performed two practice trials that had the same structure as the experimental trials, using pictures that were not used in the critical trials but were included in the booklet. The experimental session followed this practice session.

The repetition trials were not analyzed as they were only used to ensure that the distractor sentences remained in participants' memory until the picture presentation. For picture trials, the speech onset time relative to the picture onset was measured automatically using a simple amplitude threshold detection method using Matlab (version 7.13), followed by human checking. For the amplitude threshold detection method, we calibrated the amplitude threshold for each participant using the first three trials. Any trials where participants named pictures with something other than the intended target words were excluded (1.7% of experimental trials). In addition, any trials with a speech onset time of < 300 ms or > 2000 ms (0.4% of the remaining experimental trials), and any trials with onset times > 3 standard deviations away from each participant's mean were removed from the data (1.2% of the remaining data). Speech onset latency was log-transformed and submitted to statistical analyses.

A.2. Results

Participants took longer to start producing picture names in the related condition (mean = 806 ms, within-subject standard error of the mean = 4 ms) compared to the unrelated condition (mean = 784 ms, within-subject standard error of the mean = 4 ms), thus showing an interference effect typical of PWI studies. This difference was supported by a significant effect of relatedness (Relatedness: $\beta = 0.03$, $SE = 0.01$, $|t| = 2.25$, $p < .05$).

Appendix B. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2020.104183.

References

Abdel Rahman, R., & Melinger, A. (2009). Semantic context effects in language production: A swinging lexical network proposal and a review. *Language & Cognitive Processes, 24*(5), 713–734.

Allum, P. H., & Wheeldon, L. (2009). Scope of lexical access in spoken sentence production: Implications for the conceptual–syntactic interface. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(5), 1240–1255.

Anshen, F., & Aronoff, M. (1988). Producing morphologically complex words. *Linguistics, 26*(4), 641–656.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.). *The psychology of learning and motivation: Advances in research and theory* (pp. 47–89). New York: Academic Press.

Barner, D., & Bale, A. (2002). No nouns, no verbs: Psycholinguistic arguments in favor of lexical underspecification. *Lingua, 112*(10), 771–791.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01.

Bedny, M., Caramazza, A., Grossman, E., Pascual-Leone, A., & Saxe, R. (2008). Concepts are more than percepts: The case of action verbs. *Journal of Neuroscience, 28*(44), 11347–11353.

Bloem, I., van den Boogaard, S., & La Heij, W. (2004). Semantic facilitation and semantic interference in language production: Further evidence for the conceptual selection model of lexical access. *Journal of Memory and Language, 51*(2), 307–323.

Bock, J. K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review, 89*(1), 1.

Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology, 18*(3), 355–387.

Bock, J. K., & Levelt, W. (1994). Language production: Grammatical encoding. In M. Gernsbacher (Ed.). *Handbook of psycholinguistics* (pp. 945–984). San Diego, CA:

Academic Press.
Boersma, P., & Weenink, D. (2018). *Praat: Doing phonetics by computer (version 6.0.43) [computer program]*.
Caramazza, A., & Hillis, A. E. (1991). Lexical organization of nouns and verbs in the brain. *Nature, 349*(6312), 788.
Caramazza, A., Laudanna, A., & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition, 28*(3), 297–332.
Damian, M. F., & Bowers, J. S. (2003). Locus of semantic interference in picture-word interference tasks. *Psychonomic Bulletin & Review, 10*(1), 111–117.
De Smedt, K. (1996). Computational models of incremental grammatical encoding. In T. Dijkstra, & K. de Smedt (Eds.). *Computational psycholinguistics: Ai and connectionist models of human language processing* (pp. 279–307). London UK: Taylor and Francis.
Dell, G. S., Oppenheim, G. M., & Kittredge, A. K. (2008). Saying the right word at the right time: Syntagmatic and paradigmatic interference in sentence production. *Language & Cognitive Processes, 23*(4), 583–608.
Ferreira, F. (2000). Syntax in language production: An approach using tree-adjoining grammars. In L. Wheeldon (Ed.). *Aspects of language production* (pp. 291–330). .
Ferreira, V. S. (1996). Is it better to give than to donate? Syntactic flexibility in language production. *Journal of Memory and Language, 35*(5), 724–755.
Ferreira, V. S., & Humphreys, K. R. (2001). Syntactic influences on lexical and morphological processing in language production. *Journal of Memory and Language, 44*(1), 52–80.
Fromkin, V. A. (1971). The non-anomalous nature of anomalous utterances. *Language, 47*(1), 27–52.
Garrett, M. (1975). The analysis of sentence production. In G. H. Bower (Vol. Ed.), *Psychology of learning and motivation. Vol. 9. Psychology of learning and motivation* (pp. 133–177). New York: Academic Press.
Garrett, M. (1980). Levels of processing in sentence production. In B. L. Butterworth (Ed.). *Language production, volume I: Speech and talk* (pp. 177–220). London, UK: Academic Press.
Garrett, M., & Newmeyer, F. (1988). Processes in language production. *Linguistics: The Cambridge Survey, 3*, 69–96.
Griffin, Z. M. (2001). Gaze durations during speech reflect word selection and phonological encoding. *Cognition, 82*(1), B1–B14.
Iwasaki, N. (2010). Incremental sentence production: Observations from elicited speech errors in Japanese. In H. Yamashita, Y. Hirose, & J. L. Packard (Eds.). *Processing and producing head-final structures* (pp. 131–151). Dordrecht: Springer.
Janssen, N., Melinger, A., Mahon, B. Z., Finkbeiner, M., & Caramazza, A. (2010). The word class effect in the picture–word interference paradigm. *Quarterly Journal of Experimental Psychology, 63*(6), 1233–1246.
Kempen, G., & Huijbers, P. (1983). The lexicalization process in sentence production and naming: Indirect election of words. *Cognition, 14*(2), 185–209.
Konopka, A. E. (2012). Planning ahead: How recent experience with structures and words changes the scope of linguistic planning. *Journal of Memory and Language, 66*(1), 143–162.
Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 104*(2), 211–240.
Lau, E. F., Holcomb, P. J., & Kuperberg, G. R. (2013). Dissociating N400 effects of prediction from association in single-word contexts. *Journal of Cognitive Neuroscience, 25*(3), 484–502.
Laudanna, A., Badecker, W., & Caramazza, A. (1992). Processing inflectional and derivational morphology. *Journal of Memory and Language, 31*(3), 333–348.
Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*(1), 1–38.
Lukatela, G., Gligorijević, B., Kostić, A., & Turvey, M. T. (1980). Representation of inflected nouns in the internal lexicon. *Memory & Cognition, 8*(5), 415–423.
Lupker, S. J. (1979). The semantic nature of response competition in the picture-word interference task. *Memory & Cognition, 7*(6), 485–495.
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*(3), 503.
Marslen-Wilson, W., Tyler, L. K., Waksler, R., & Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychological Review, 101*(1), 3.
Melinger, A., & Koenig, J.-P. (2007). Part-of-speech persistence: The influence of part-of-speech information on lexical processes. *Journal of Memory and Language, 56*(4), 472–489.
Meyer, A. S. (1996). Lexical access in phrase and sentence production: Results from picture–word interference experiments. *Journal of Memory and Language, 35*(4), 477–496.
Momma, S., & Phillips, C. (2018). The relationship between parsing and generation. *Annual Review of Linguistics, 4*, 233–254.
Momma, S., Slevc, L. R., & Phillips, C. (2016). The timing of verb selection in Japanese sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(5), 813–824.
Momma, S., Slevc, L. R., & Phillips, C. (2018). Unaccusativity in sentence production. *Linguistic Inquiry, 49*(1), 181–194.
Moseley, R. L., & Pulvermüller, F. (2014). Nouns, verbs, objects, actions, and abstractions: Local fmri activity indexes semantics, not lexical categories. *Brain and Language, 132*, 28–42.
Nooteboom, S. G. (1980). The tongue slips into patterns. In G. Sciaron, A. van Essen, & A. Van Raad (Eds.). *Leyden studies in linguistics and phonetics* (pp. 114–132). The Hague, Netherlands: Mouton.
Pechmann, T., Garrett, M., & Zerbst, D. (2004). The time course of recovery for grammatical category information during lexical processing for syntactic construction. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(3), 723–728.
Pechmann, T., & Zerbst, D. (2002). The activation of word class information during speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(1), 233–243.
Pfau, R. (2009). *Grammar as processor: A distributed morphology account of spontaneous speech errors.* Amsterdam: John Benjamins Publishing.
Potter, M. C., & Lombardi, L. (1990). Regeneration in the short-term recall of sentences. *Journal of Memory and Language, 29*(6), 633.
R Core Team (2019). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. (URL) https://www.R-project.org/.
Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition, 42*(1–3), 107–142.
Sachs, J. S. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Perception & Psychophysics, 2*(9), 437–442.
Schnur, T. T., Costa, A., & Caramazza, A. (2006). Planning at the phonological level during sentence production. *Journal of Psycholinguistic Research, 35*(2), 189–213.
Schriefers, H., Meyer, A. S., & Levelt, W. J. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language, 29*(1), 86–102.
Schriefers, H., Teruel, E., & Meinshausen, R.-M. (1998). Producing simple sentences: Results from picture–word interference experiments. *Journal of Memory and Language, 39*(4), 609–632.
Smith, M., & Wheeldon, L. (1999). High level processing scope in spoken sentence production. *Cognition, 73*(3), 205–246.
Stemberger, J. P. (1985). An interactive activation model of language production. In A. W. Ellis (Vol. Ed.), *Progress in the psychology of language. Vol. 1. Progress in the psychology of language* (pp. 143–186). London: Erlbaum.
Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., … others (2004). A new on-line resource for psycholinguistic studies. Journal of Memory and Language, 51 (2), 247–250.
Vigliocco, G., Vinson, D. P., Druks, J., Barber, H., & Cappa, S. F. (2011). Nouns and verbs in the brain: a review of behavioural, electrophysiological, neuropsychological and imaging studies. *Neuroscience & Biobehavioral Reviews, 35*(3), 407–426.
Vigliocco, G., Vinson, D. P., Lewis, W., & Garrett, M. (2004). Representing the meanings of object and action words: The featural and unitary semantic space hypothesis. *Cognitive Psychology, 48*(4), 422–488.
Vigliocco, G., Vinson, D. P., & Siri, S. (2005). Semantic similarity and grammatical class in naming actions. *Cognition, 94*(3), B91–B100.
Wagner, V., Jescheniak, J. D., & Schriefers, H. (2010). On the flexibility of grammatical advance planning during sentence production: Effects of cognitive load on multiple lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(2), 423–444.
Wardlow-Lane, L., & Ferreira, V. S. (2010). Abstract syntax in sentence production: Evidence from stem-exchange errors. *Journal of Memory and Language, 62*(2), 151–165.
Yuan, J., & Liberman, M. (2008). Speaker identification on the SCOTUS corpus. *Journal of the Acoustical Society of America, 123*(5), 3878.
Zwitserlood, P., Bölte, J., & Dohmes, P. (2000). Morphological effects on speech production: Evidence from picture naming. *Language & Cognitive Processes, 15*(4–5), 563–591.
Zwitserlood, P., Bölte, J., & Dohmes, P. (2002). Where and how morphologically complex words interplay with naming pictures. *Brain and Language, 81*(1–3), 358–367.