Syntactic category constrains lexical competition in speaking

Shota Momma
University of California, San Diego

Julia Buffinton
University of Maryland, College Park

L. Robert Slevc
University of Maryland, College Park

Colin Phillips
University of Maryland, College Park

Author Note

Shota Momma, Department of Psychology, University of California, San Diego
This research was supported in part by the grants from National Science Foundation grant BCS-0848554 to Colin Phillips and National Institutes of Health grant R01-HD051030 to Victor Ferreira.

Correspondence concerning this article should be addressed to Shota Momma, Department of Psychology, University of California, San Diego, 9500 Gilman Drive. La Jolla, CA, 92093. Contact: smomma@ucsd.edu

Abstract

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 articulate sentences mostly fluently and relatively effortlessly. The current article provides experimental evidence that syntactic category plays a key role in limiting competition during lexical access in speaking. We introduce a novel sentence-picture interference SPI paradigm, and show that nouns do not compete with verbs and verbs do not compete with nouns in sentence production. Words that are conceptually and phonologically identical, such as *running* (noun) and *running* (verb), lead to interference only when they match in syntactic category. Based on this finding, we argue that lexical competition in production is limited by syntactic category. We discuss the potential underlying mechanism and how it may help us to speak relatively fluently.

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

Syntactic category constrains lexical competition in speaking

Introduction

Saying a word feels easy, but it requires accessing an appropriate representation of the word from among tens of thousands of words in speakers' mental dictionaries (Nagy, Herman, & Anderson, 1985; Pinker, 1994), many of which are similar to each other. Just like other types of memory access, this process is subject to similarity-based retrieval interference from non-target items (Levelt, Roelofs, & Meyer, 1999; Roelofs, 1992). Lexical access requires overcoming competition from these similar words, and competition is likely even greater when saying a sentence because speakers must access multiple words in a rapid, specifically-ordered sequence, while each word accessed creates an additional source of interference for the others (Dell, Oppenheim, & Kittredge, 2008). Yet healthy adult native speakers say sentences mostly fluently and relatively effortlessly. How is our lexical memory organized and accessed in such a way that allows us to speak so fluently and effortlessly? The current article provides experimental evidence 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 do not compete with verbs 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 discuss the potential underlying mechanism and how it may enable us to speak relatively fluently.

Competition in lexical access

In single word production research it is widely assumed that items in lexical memory are selected competitively (Levelt et al., 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. For instance, when accessing the lemma for *cat*, activation of a conceptually similar lemma such as *dog* can interfere with the 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 relations, specifically those 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 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; Nootebaum, 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. Dell et al. (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 the competition between words of different categories. The role of the sentence context in Dell and colleagues' model fits with evidence from PWI experiments pitting 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, partly because the speech error evidence and experimental evidence diverge from each other. For example, Vigliocco et al. (2005), discussed above, is cited as converging with the speech error pattern (Dell et al., 2008), but the convergence with the speech error evidence is only partial. Specifically, Vigliocco and colleagues did find cross-category interference effects, albeit less pronounced than within-category interference effects. Word exchange 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 27ms 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, and Zerbst (2004); Pechmann and 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, & Caramazza, 2010), and, in any case, they do not show that syntactic category has a competition-limiting function. Instead, they simply show that there is additional interference from syntactic category overlap.

What, then, explains the disconnect between the speech error record and the reliable (even if small) 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*; *cry*) before being constrained to a nominal form via combination with *il*. A second possibility is that the manipulation of syntactic category in Vigliocco et al.'s Vigliocco et al. (2005) experiments was confounded with morphological overlap. Specifically, Italian has three classes of verbs that end with either *-are*, *-ere*, and *-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 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 if 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.

Implications of the syntactic category constraint

If syntactic category limits competition in lexical access, speakers must be able to both (i) project the syntactic category of upcoming words ahead of retrieval and (ii) restrict retrieval to candidate words that match the projected syntactic category. Thus, examining the extent to which the syntactic category constraint applies is relevant to the study of sentence production more generally, and also to the study of lexical memory used in speaking. Regarding (i), if syntactic category constrains lexical retrieval, then the sentence production mechanism must represent information about the syntactic category of an upcoming word prior to actually selecting that word. Thus, evidence for the syntactic category constraint in lexical access is also evidence that the production mechanism is capable of representing structure before content. Hence, examining the syntactic category constraint can usefully inform general models of sentence production because it informs how structural processes constrain lexical processes.

Regarding (ii), the syntactic category constraint implies that the target of lexical access (i.e., a to-be-retrieved lexical item) carries syntactic category information that can be matched with the projected syntactic category representation. The strongest version of this claim is that even a morphologically complex word is stored together with its syntactic category information. For example, *categorization* is a morphologically complex noun that is internally a verb, which in turn is derived from a noun (see the next section for more detail). In theory, lexical access could target the initial representation, intermediate representation, or only the final representation of this derivational process. However, in order for the syntactic category constraint to limit competition, the projected syntactic category (represented prior to lexical access) needs to match the target of lexical access in syntactic category. Of course, the syntactic category constraint might only be effective for morphologically simple words, but that would render the constraint not very useful in most situations. Indeed, some linguistic theories (see the next section) postulate that every single word is morphologically complex and is built from morphemes on the fly in processing. Thus, the effectiveness of the syntactic category constraint depends on the the representational format of lexical memory: If words are not stored together with their syntactic category information or derived from category-changing derivational morphological processes every time they are spoken, the syntactic category constraint should not be able to constrain lexical access. Thus the existence of the syntactic category constraint has implications for the status of syntactic categories in lexical memory.

The status of syntactic categories in lexical memory

How is lexical memory organized and accessed? A leading theory in psycholinguistics is that speakers store and access the representation of a whole word, possibly to the exclusion of inflectional morphology, with its syntactic category information specified as a part of each entry (Fromkin, 1971; Levelt et al., 1999). One source of evidence for this claim comes from neuropsychological research that shows selective deficits in either noun production or verb production (Berndt, Mitchum, Haendiges, & Sandson, 1997; Hillis & Caramazza, 1991). The syntactic category constraint in speech errors also lends support to the idea that whole words are stored together with their syntactic category information. Though alternative interpretations are possible (Barner & Bale, 2002; Pfau, 2009), a widely adopted interpretation of these data is that lexical memory contains whole word representations and is organized according to syntactic category. Under the strong version of this view, even complex words that are derived via productive morphological rules, like *singing* as a noun (as in *her skillful singing*, are stored based on the category of their final representation, and so a single retrieval process is sufficient to arrive at the final representation of the noun *singing*. Under this account, lexical memory contains morphologically complex entries with their syntactic category specified.

An alternative possibility is that derived complex words are not stored as a whole, and are underspecified for their syntactic category information. This view is not widely adopted in production research or in psycholinguistics more generally, though with some exceptions (Barner & Bale, 2002; Garrett, 1975; Pfau, 2009). However, it is nevertheless a reasonable possibility that would presumably also be more economical in terms of memory storage. This position is appealing not only because it assumes more economical long-term memory representations, but also because the syntactic category of words can flexibly and productively change in many languages, including English. This position is in line with a linguistic theory called Distributed Morphology, which posits that the morphological roots of words do not contain syntactic category information (Halle & Marantz, 1993; Harley & Noyer, 1999; Marantz, 1997). Distributed Morphology (often) postulates that syntactic categories are supplied by an unpronounced functional element that is distinct from roots (henceforth the categorizer), and hence claims that roots themselves are not specified for their syntactic category. In other words, every word consists of at least two morphological atoms, a root and a categorizer, and often more. Distributed Morphology was not designed as a theory of processing. Thus, the claim of the distributed morphology about syntactic category does not prevent the assumption that some complex words are stored in memory, and hence can be targeted by retrieval processes. However, the core assumption about syntactic category may apply transparently to acquisition and processing (Barner & Bale, 2002; Pfau, 2009). Such a view suggests that items in lexical memory are not words, but rather morphological atoms, and that the retrieval process targets morphemes rather than whole words. For instance, when speaking the word *categorization*, speakers must retrieve a category-less root *category* (often denoted as $\sqrt{}$category), then combine it with the independently retrieved morpheme *-ize* to form a verb, and then finally combine the product with another independently retrieved morpheme *-tion* to reach its eventual status as a noun.

Because these two processing hypotheses make divergent claims about the relative timing at which a word gains its eventual syntactic category status, they make diverging predictions about the circumstances where syntactic category can limit retrieval interference. According to the hypothesis that whole words are stored and specified for their eventual syntactic category, syntactic category information is a feature that can be directly used to constrain memory access. In other words, syntactic category information can impact retrieval processes from the earliest stage of retrieval. In contrast, according to the hypothesis that syntactic category is not specified for each memory entry, syntactic category logically cannot limit retrieval interference, because the eventual syntactic category is only available after all retrieval processes are finished.

In the current study we test these predictions using nominal gerunds, which appear in nominal contexts but preserve a defining property of the verb, since they are argument-taking (Abney, 1987; Alexiadou, 2001; Chomsky, 1970; Grimshaw, 1990). For instance, *singing* as a noun, as in *Mary's skillful singing of a song*, takes a subject argument *Mary* and an object argument *a song*. For an obligatorily transitive verb (unlike *sing*, which is optionally transitive), the object argument is also obligatory, e.g., *John's destroying (of) the city* is acceptable but*John's destroying* is not. This inheritance of argument structure properties is not observable in nominals like destruction, which only optionally takes the argument despite the root 'verb' being obligatorily transitive. For this reason (among other reasons), nominal gerunds are often analyzed as internally containing verbs, e.g., the nominal *singing* is a derivative of the verb *sing* (Abney, 1987; Kratzer, 1996). Under the hypothesis that each entry in our lexical memory is specified for its eventual syntactic category, speakers do not (have to) represent a category-less root $\sqrt{}$whistle or verb *whistle* when saying *whistling* as a noun. Of course, this hypothesis does not prevent the category less roots and intermediate morpheme complex to be represented in memory in addition. Combined with the claim that syntactic category limits retrieval interference (see Section 1.1 above), this hypothesis predicts that a related word like *singing* as a noun should not interfere with the production of a conceptually related verb like *whistling.* In contrast, the hypothesis that entries in lexical memory are underspecified for their syntactic category posits that *whistling* as a noun is computed from the category-less root $\sqrt{}$whistle, and, in the course of deriving the final noun representation, there is a point at which *whistle* is represented as a verb. Under this hypothesis, a conceptually related word like *singing* as a noun should interfere with *whistling* as a verb, because *singing* and *whistling* are conceptually similar and their eventual syntactic category status is not assigned until retrieval processes are complete. Although the current study does not directly examine the mental representation of syntactic category, it does address this issue indirectly. This is because different assumptions about the representation of syntactic category predict different interference profiles between morphologically complex words. Thus the results of the current study are relevant to these larger questions about the nature of syntactic representations in lexical memory.

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. The predicted pattern of interference relates both to the existence of a syntactic category constraint and to the issue of whether the targets of lexical retrieval are complex representations specified for their syntactic category or roots that are underspecified for their syntactic category. Together, these questions yield four possible hypotheses, summarized in Table 1 along with their predictions about the pattern of interference effects from conceptually related distractors.

To test these hypotheses, there are some experimental challenges that need to be overcome. The above discussion of Vigliocco et al. (2005) 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 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 in a fleeting amount of time.

To test the above hypotheses using a more appropriate experimental paradigm that does not have the shortcomings of PWI, we introduce a novel experimental task that we named the sentence-picture interference (SPI) task, illustrated in Figure 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 perfectly 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. 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. The basic task structure is illustrated in Figure 1.

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 speakers store *singing* as a whole, either as a verb or as a noun. 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 three experiments that we report below all use the SPI task, and they are closely related to each other. The first experiment simply tests whether the SPI task is able to elicit a semantic interference effect just like PWI in a simple object naming context. The second and third experiments test the critical question of the current article: whether words from two different syntactic categories compete with each other.

Experiment 1

Experiment 1 tests the effectiveness of the sentence-picture interference (SPI) task. To our knowledge, this task 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, in Experiment 1 we aimed to test whether the SPI task is sensitive to the semantic interference effect that is usually seen in the PWI task. If successful, we can also use the interference effect from Experiment 1 as a baseline for the presence or absence of interference in the subsequent experiments.

Methods

Participants. Twenty-four University of Maryland undergraduate students participated in Experiment 1 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.

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 and Dumais 1997). The related distractor sentences were re-paired with another target to create unrelated target-distractor pairs. That is, the set 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 Figure 2. Because the set of related and unrelated distractor sentences were identical, differences in sentence complexity, plural/singular differences, etc., cannot explain the 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).

Procedure and Analysis. Each experimental trial was structured as follows. First, the participants saw a fixation cross at the center of the screen for 300ms. Following a 200ms 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 300ms. Following a 200ms black screen, a picture stimulus appeared on the screen for 5000ms on 50 percent of the trials. In that case, participants responded by speaking the target word that corresponded to the picture stimulus. In the other 50 percent 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 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 they saw is available at https://shotam.github.io/CategorySPI/Pictures.pdf. After this familiarization session, the structure of each trial (illustrated in Figure 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 (1.6 percent of experimental trials). In addition, any trials with a speech onset time of less than 300ms or more than 2000ms (0.4 percent of the remaining experimental trials), and any trials with onset times more than 3 standard deviations away from each participant's mean were removed from the data (1.2 percent of the remaining data). Speech onset latency was log-transformed and submitted to statistical analyses. All the statistical analyses were conducted using R (R Core Team, 2017), and the lme4 package (Bates, Mächler, Bolker, & Walker, 2015). For the mixed effects model analysis, all the categorical experimental factors were sum coded (i.e., 0.5, -0.5). The random effects structure was maximal in the sense of Barr, Levy, Scheepers, and Tily (2013).

Results

Participants took longer to start producing picture names in the related condition (mean = 806 ms, within-subject standard error of the mean = 4.28 ms) compared to the unrelated condition (mean = 784 ms, within-subject standard error of the mean = 4.28 ms), thus showing an interference effect typical of PWI studies. This difference was supported by a significant effect of relatedness in the mixed effects model reported in Table 4 (Relatedness: $\beta = 0.03, SE = 0.01, |t| = 2.21, p < .05$). 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 trails 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 variable, to determine whether trial order or the interaction term should be included. Including the centered trial order improved the model fit ($\chi^2(1) = 22.68, p < .01$). The interaction between the centered trial order and relatedness did not improve the model fit ($\chi^2(1) = 2.09, p > .1$), and thus was not included in the final model.

Discussion

Experiment 1 showed that semantically related distractors can cause interference effects in the SPI task, just like in the traditional PWI task. Thus, the SPI task is a suitable task for assessing the presence or absence of semantic interference effects in lexical access.

One critical difference between the SPI task and the PWI task merits discussion. Namely, the relative timing at which the distractor is presented to the participants is different. 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 Bloem, van den Boogaard, and La Heij (2004) in word-translation interference tasks. 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, the distractor presentation occurs well before the 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.

One potential reason is that, unlike in typical PWI experiments or in the translation interference experiments by Bloem et al. (2004), we forced speakers to memorize the distractors. 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). Hence, despite the critical difference in the results, we argue that the SPI task is consistent with the model of distractor effects built to explain the results of PWI experiments.

Experiment 2

Experiment 2 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 2 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. [1] This means that speech onset latency alone is not a suitable measure for assessing the presence or absence of interference effects in a sentence production task.

Method

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

[1]In both Experiments 2 and 3, we also conducted analyses on the estimates of speech onset latency obtained from the simple amplitude threshold detection method used in Experiment 1. The results of the statistical significance patterns for both Experiment 2 and 3 were the same.

[2]In Experiment 2 we choose to test 48 participants rather than 24, in order to increase the chance of detecting a non-crossover interaction effect.

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. Just like in Experiment 1, 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 Figure 2. Unlike in Experiment 1, each participant never saw the same sentence templates more than 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 2 (and Experiment 3) is available at https://shotam.github.io/CategorySPI/StimList.csv.

Procedure and Analysis. The basic experimental procedures were the same as in Experiment 1. However, the stimulus presentation timing was changed so that the fixation cross appeared for 500ms, followed by a 300ms blank screen. Additionally, the offset of distractor sentences and the onset of the picture stimulus/repeat prompt were separated by a 1000ms black screen. These changes were made in part due to feedback from the test participants in Experiment 1, and also because we thought that seeing category-ambiguous critical words like *whistling* might transiently activate both nominal and verbal representations. If both representations are indeed transiently activated, it would prevent us from investigating the critical question of whether between-category lexical representations compete with each other. This was not a problem in Experiment

1 because all critical words were clearly nouns. Hence, we introduced this interval in Experiments 2 and 3, but not in Experiment 1. For statistical analyses, any responses that deviated from the target sentences were excluded (4 percent of all experimental trials). Trials with silent pauses were preserved, as pauses might reflect semantic interference effects. Any trials with a speech onset time of less than 300ms or more than 5000ms (0.5 percent of the remaining 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 percent 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 measured verb production latency using a text-to-speech alignment algorithm (Penn Forced Aligner; Yuan and Liberman 2008).The resulting values of the first and last ten trials of the first and last three participants (120 trials in total) were evaluated for their interclass correlation (Shrout & Fleiss, 1979). This revealed that the consistency between the automatic alignment method and the human-coding was *excellent* (ICC = 0.90), according the criteria set by Cicchetti (1994). Previous studies have shown that measuring production timing of sentence-internal elements using the forced alignment technique is effective for detecting semantic interference effects that occur later than speech onset (Momma et al., 2018). All the pairwise comparisons were based on planned subset analyses. The interaction terms involving trial order were removed from the model because they did not significantly improve the model fit. The random effects structure was maximal in the sense of Barr et al. (2013).

Results

As can be seen in Table 2 speakers took around 40 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 5, this contrast was supported by an interaction between relatedness and distractor category ($p < .05$)

as well as by planned comparisons showing the significant effect of relatedness for verb distractors ($\beta = 0.03$, $SE = 0.01$, $|t| = 3.01$, $p < .01$) but not for noun distractors ($\beta = 0.01$, $SE = 0.01$, $|t| = 0.64$, $p > .5$).

Discussion

The results show that interference from conceptually similar words in distractor sentences slowed production of verbs selectively, 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 the consistent retrieval of an intermediate representation when accessing *singing* and *whistling*, because models that assume the retrieval of an intermediate representation should predict an interference effect between *singing* and *whistling* regardless of their syntactic category. If accessing *whistling* involves an intermediate step of accessing the category-less root $\sqrt{}$whistle, interference effects are expected between the distractor and target, even when the distractor words are from a different syntactic category. This was not the case in the current experiment. Thus, the current results are more compatible with a model in which lexical representations are stored and accessed with their syntactic category status.

However, there were some unavoidable confounds between the verb and noun distractor conditions in Experiment 2. 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. Indeed, there were numerical (non-significant) differences (19 ms) between the unrelated conditions in the noun distractor vs. verb distractor conditions, perhaps because of these confounds. It is possible that this difference somehow interacted with the semantic interference effects. Thus, Experiment 3 tested whether the reverse pattern of interference effects can be obtained when speakers produce nouns instead of verbs.

Experiment 3

Experiment 3 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 2. 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 Figure 3 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...

After a couple of practice trials, these instructions reliably elicited sentences containing gerundive nominals, e.g., *her whistling is red.* Using this method, Experiment 3 tested whether noun production was selectively interfered with by noun distractors. In Experiment 3 we measured noun production latency rather than speech onset latency, for the same reason we measured verb production latency in Experiment 2.

Method

Participant. Twenty-four University of Maryland undergraduate students participated in Experiment 3 for course credit. Informed consent was obtained from all the participants prior to the experiment. None had participated in Experiment 1 or 2.

Materials. An example picture stimulus is presented in Figure 3. The pictures (except the colored square) and distractor sentences were identical to those used in Experiment 2.

Procedure and Analysis. The experimental procedures and analyses were identical to Experiment 2, except that participants were given different instructions for the picture description component of the task, as described above.

Results. As can be seen in Table 3, speakers took about 45 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 -7 ms difference). This was supported by an interaction between relatedness and distractor category (Table 6) as well as by planned comparisons showing a significant effect of relatedness for noun distractors ($\beta = 0.03$, $SE = 0.02$, $|t| = 2.40, p < .05$) but not for verb distractors ($\beta = 0.00$, $SE = 0.01$, $|t| = 0.23, p > .8$). As in Experiments 1 and 2, 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 they are removed from the final model. The random effects structure was maximal in the sense of Barr et al. (2013).

Discussion

In Experiment 3, 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 2 are more consistent with models that do not assume retrieval of intermediate representations in accessing complex gerundive nominals. This is because the syntactic category constraint should only be effective when the projected syntactic category matches the syntactic category of the retrieval target. If accessing *whistling* as a noun involves an intermediate step of accessing the category-less root $\sqrt{}$whistle, or the verb *whistle*, 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 3 or in Experiment 2. 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 or intermediate representations of derivational processes are involved in word retrieval processes.

The combination of verb-specific and noun-specific interference found in Experiments 2 and 3 further shows that the potential confounding differences in naturalness, length, and complexity between distractor conditions in Experiment 2 were unlikely to be the cause of the selective interference pattern. This is because Experiment 3 used the exact same set of distractors (and pairing with target pictures), thus had exactly the same potential confounds, yet showed the opposite pattern of results (see Figure 4 below).

General Discussion

The current experiments yielded three main results. Experiment 1 showed that the SPI task is suitable for eliciting semantic interference effects. Experiment 2 demonstrated that the semantic interference effect in verb production is observable only when the distractors are also verbs. Experiment 3 showed that the semantic inference effect in noun production is observable only when the distractors are also nouns. The pattern of interference in Experiments 2 and 3 is summarized in Figure 4. This interference pattern suggests that competition in lexical access is restricted to words from the same syntactic category. The within-category restriction of competition is readily explained if the syntactic category of a whole word is stored and accessed together with its syntactic category information.

Mechanisms

The current experiments showed that semantic interference effects in the SPI task are restricted to words that share the same syntactic category. A plausible interpretation of this pattern is that words are stored and accessed together with their syntactic category information, so words in a different syntactic category do not compete for selection. Here we discuss some potential cognitive mechanisms that may underlie this selective nature of lexical competition.

One mechanism that might explain the selective nature of lexical competition is the syntactic traffic cop mechanism described by Dell et al. (2008). In the 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 nodes with positive weights (e.g., *dog*). In contrast, lexical nodes that are connected with negative weights (e.g., *bark*) are inhibited. This minimizes the effect of competition from inhibited lexical nodes, that is, words in different syntactic categories than the target.

The pattern of interference effects predicted by the syntactic traffic cop mechanism depends on the activation function of the network. If the activation function is linear, the model predicts an additive effect of syntactic category overlap and semantic overlap. In other words, the model predicts semantic interference effects between similar words regardless of their syntactic category, but the effect should be stronger between words from the same category. As Dell et al. discussed, the results of Vigliocco et al. (2005) follow this pattern (but see an alternative interpretation discussed in the Introduction). On the other hand, if the activation function is non-linear (logistic/sigmoidal), it may be possible to configure the connection weights appropriately so that only the nodes that share the same syntactic category and semantic features will be excited and compete with the target selection. The current results suggest that the activation function of the syntactic traffic cop mechanism should be non-linear, because we do not see any semantic interference effects between words of different categories. In this sense, the current data may be seen as constraining the parameter space of Dell's account of syntactic category effects.

An alternative potential mechanism is essentially analogous to the bin model of lexical recognition described by K. Forster (1996); K. I. Forster (1992). In Forster's model, lexical memory is first partitioned into multiple *bins*, and bins are searched in parallel (but serially within each bin). As a result, words in different bins do not affect the time it takes to access the target words. Although Forster's model is a model of lexical recognition, it is possible that this binning mechanism is shared with the model of lexical access in speaking. With a couple of additional assumptions, this architecture might be suitable for capturing the current results. First, the model needs to assume that search within a single bin is subject to competition based on conceptual similarity. Second, the model needs to assume that lexical items are partitioned into different bins according to their syntactic category. With these assumptions, this model might capture the selective interference effects in lexical selection observed here.

However, there is an important conceptual challenge to both Dell's syntactic traffic cop mechanism and the bin model: the problem of memory economy and the problem of learning. In order to capture non-competition across syntactic categories (as shown in the current studies), Dell's approach has to assume that verbs and nouns are represented separately by different nodes. For example, the model has to have one node representing *whistling* as a noun and another node representing *whistling* as a verb. Otherwise, the model cannot have two different connection weights of opposite polarity. This duplication presumably has to happen for every verb in English, because *-ing* gerundive morphology is completely productive in English. This creates a memory economy problem, because every verb needs to also be redundantly represented as a noun. This problem of memory economy is accompanied by a potentially more serious learning problem. Learners need to somehow ensure that relevant properties of a verbal representation (e.g., collocational information, semantics, etc) transfer to the corresponding nominal representation (and vice versa), and they need to also ensure that the two representations do not diverge in their syntactic and semantic properties. Similarly, in the bin model, each bin needs to contain *whistling*, one as a noun and the other as a verb. This creates a similar redundancy as in Dell's approach. Thus, essentially the same problem of memory economy and learning arises, whichever the mechanism one chooses to adopt.

Having said this, it is currently unclear if our brain stores memories in a way that maximizes economy in terms of memory space. The strategy to minimize memory space comes with the potential cost of needing to compute complex representations each time such representations are involved in online production. Thus, the "economical" storage of lexical memory may not be necessarily economical in terms of real-time processing cost, for the very reason we show here that syntactic category is helpful in limiting retrieval interference. Additionally, it is possible that selective competition only applies to the verbs that frequently undergo nominalization. If that is the case, then lexical memory only needs to duplicate verbs that are frequently used in a nominal context in everyday life. This may reduce the memory economy problem, while explaining the selective nature of competition in the current study. It is currently unclear if the current results extend to gerundive nominals that are rarely or never naturally used, so this possibility cannot be eliminated by the current results. Regarding the learning problem, in order to assess the plausibility of the mechanisms presented here, there needs to be a theory of how learning transfers between separate representations, and how the learning mechanism ensure that two related representations do not diverge in terms of their semantic and syntactic properties. These are questions for the future studies.

Implications for models of speaking

The current finding also informs processing models of sentence production more generally. Logically speaking, in order for syntactic category to constrain lexical access, syntactic category must be represented prior to the point of lexical access. This contrasts with well known lexicalist models of production (e.g., Bock and Levelt 1994), in which syntactic categories are projected from words in a *bottom-up* fashion. Under the account proposed here, speakers are able to project 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, nor that the syntactic structure of an entire clause needs to be planned in advance (e.g., F. Ferreira 2000; Garrett and Newmeyer 1988). 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, at least when speakers can be certain about the category of an upcoming word as in the current experiments, they use the category information to constrain lexical access. This *top-down* method of building syntactic structure may be useful for speakers to minimize interference.

Furthermore, 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, Teruel, & Meinshausen, 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 confirm to the selectional requirement of verbs, without increasing memory cost (Momma, Slevc, & Phillips, 2016; Momma et al., 2018). This may, therefore, be why speaking can be both relatively fluent and grammatically robust.

The status of syntactic category in lexical memory

The current results also suggest that the syntactic category of a word is already available as the word is retrieved. This in turn suggests that words are stored in memory, together with their syntactic category information. Of course, as acknowledged above, this does not mean that every single complex word is stored in lexical memory and accessed via a single retrieval process. Indeed, this 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. Instead, the current results suggest that speakers are capable of storing and accessing memories of complex words that are specified for their syntactic category in addition to smaller parts (morphological atoms). While this may seem costly in terms of long-term memory space, it is likely beneficial for limiting interference in lexical access, and thus is helpful for speaking fluently. If memory of only morphological atoms were available, it is unclear how speakers could use the projected syntactic category to constrain lexical access. For example, speakers may be certain that an upcoming word is a noun after a determiner, but this is not helpful in reducing retrieval interference if the retrieval can only target a category-less root or verb stem that ends up being a noun only after derivational morphological process. Thus, we argue that speakers can (i) actively project the syntactic category of an upcoming word, and (ii) use it to (selectively) access a memory entry whose syntactic category matches the projected category.

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 a whole word is used in lexical access to limit competition even in non-erroneous speech, and thus show that the same mechanism can explain latency data in experimental settings (as shown here) and speech error data in naturalistic settings. The current results also allow us to draw a more radical conclusion than the speech error data on whole-word exchange errors. Namely, the current results suggest that even morphologically complex words that can be derived from entirely productive nominalization rules are nevertheless stored in and accessed as a whole from lexical memory during speaking. This conclusion was hard to draw solely based on the speech error data, as errors are sparse and contain only occasional instances of whole-word exchanges between morphologically complex words derivable via productive morphological rules.

One interesting type of errors that are seemingly in conflict with the current claim are 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 V. S. Ferreira and 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 syntactic category and phrasal membership of words are determined, and in which whole-word exchange errors occur. As a result, stranding errors are less sensitive to the syntactic category constraint than whole-word exchange errors, because syntactic category information is not relevant to the positional level of processing. Thus, 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 and Levelt 1994).

Alternative interpretations

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 Experiment 1. 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 retrieved word is integrated into the overall representation of a sentence. Under this interpretation, the post-lexical integration process is somehow constrained by syntactic category in such a way that the integration process is easier when distractor words have a different syntactic category. We certainly acknowledge that the 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 as it does not explain why a similar interference effect was also observed in the single word naming task in Experiment 1 (just like in PWI tasks involving single-word production). Furthermore, we believe that the current experiments offer useful results even under this alternative interpretation. First, they show that syntactic-category somehow limits post-lexical competition. Second, they show that this integration process must target 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.

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.

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 can speak fluently and grammatically at the same time, at least most of the time.

References

Abney, S. P. (1987). *The english noun phrase in its sentential aspect* (Unpublished doctoral dissertation). Massachusetts Institute of Technology.

Alexiadou, A. (2001). *Functional structure in nominals: nominalization and ergativity* (Vol. 42). Amsterdam: John Benjamins Publishing.

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.

Allum, P. H., & Wheeldon, L. R. (2007). Planning scope in spoken sentence production: the role of grammatical units. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(4), 791–810.

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. doi: 10.18637/jss.v067.i01

Berndt, R. S., Mitchum, C. C., Haendiges, A. N., & Sandson, J. (1997). Verb retrieval in aphasia. 1. characterizing single word impairments. *Brain and language*, *56*(1), 68–106.

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., & Levelt, W. (1994). *In m. gernsbacher (ed.), language production: Grammatical encoding (pp. 945–984). handbook of psycholinguistics.* San Diego, CA: Academic Press.

Chomsky, N. (1970). Remarks on nominalization. In P. R. R. Jacobs (Ed.), *Readings in english transformational grammar* (pp. 184–221). Waltham MA: Ginn.

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological assessment*, *6*(4), 284.

Damian, M. F., & Bowers, J. S. (2003). Locus of semantic interference in picture-word interference tasks. *Psychonomic Bulletin & Review*, *10*(1), 111–117.

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 and cognitive processes*, *23*(4), 583–608.

De Smedt, K. (1996). Computational models of incremental grammatical encoding. In K. d. S. A. Dijkstra (Ed.), *Computational psycholinguistics: Ai and connectionist models of human language processing* (pp. 279–307). London UK: Taylor and Francis.

Ferreira, F. (2000). Syntax in language production: An approach using tree-adjoining grammars. *Aspects of language production*, 291–330.

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.

Forster, K. (1996). Accessing the mental lexicon. In E. R.J.Wales (Ed.), *New approaches to language mechanisms* (pp. 257–287). Amsterdam: North-Holland.

Forster, K. I. (1992). Memory-addressing mechanisms and lexical access. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (Vol. 94, pp. 413–434). Elsevier.

Fromkin, V. A. (1971). The non-anomalous nature of anomalous utterances. *Language*, 27–52.

Garrett, M. (1975). The analysis of sentence production. In G. H. Bower (Ed.), *Psychology of learning and motivation* (Vol. 9, 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.

Grimshaw, J. (1990). *Argument structure.* the MIT Press.

Halle, M., & Marantz, A. (1993). Distributed morphology and the pieces of inflection. In K. Hale & S. J. Keyser (Eds.), *The view from building 20: essays in honor of sylvain bromberger.* Cambridge, MA: MIT Press.

Harley, H., & Noyer, R. (1999). Distributed morphology. *Glot international*, *4*(4), 3–9.

Hillis, A. E., & Caramazza, A. (1991). Category-specific naming and comprehension impairment: A double dissociation. *Brain*, *114*(5), 2081–2094.

Iwasaki, N. (2010). Incremental sentence production: Observations from elicited speech errors in japanese. In *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.

Kratzer, A. (1996). Severing the external argument from its verb. In J. Rooryck & L. Zaring (Eds.), *Phrase structure and the lexicon* (pp. 109–137). Dordrecht: Springer.

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.

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.

Lupker, S. J. (1979). The semantic nature of response competition in the picture-word interference task. *Memory & Cognition*, *7*(6), 485–495.

Marantz, A. (1997). No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *University of Pennsylvania working papers in linguistics*, *4*(2), 201–225.

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., 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.

Nagy, W. E., Herman, P. A., & Anderson, R. C. (1985). Learning words from context. *Reading research quarterly*, 233–253.

Nootebaum, 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). Hague, Netherland: 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* (Vol. 137). John Benjamins Publishing.

Pinker, S. (1994). *Language instinct.* London UK: Penguin Book.

R Core Team. (2017). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `http://www.R-project.org/`

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, *42*(1-3), 107–142.

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.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological bulletin*, *86*(2), 420–428.

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 (Ed.), *Progress in the psychology of language* (Vol. 1). 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., 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), 423-44.

Yuan, J., & Liberman, M. (2008). Speaker identification on the scotus corpus. *Journal of the Acoustical Society of America*, *123*(5), 3878.

Does syntactic category limit competition?	What is the target of retrieval process?	Prediction
Yes	Complex word (*whistling* [v.])	Within category interference only
No	Complex word (*whistling* [v.])	Between and within category interference
Yes	Root ($\sqrt{}$*whistle*)	Between and within category interference
No	Root ($\sqrt{}$*whistle*)	Between and within category interference

Table 1

The summary of four hypotheses and their predictions, using production of the verb whistling as an example.

Relatedness	DistractorType	mean	se
Related	Noun	1302.36	9.66
Related	Verb	1326.12	7.39
Unrelated	Noun	1300.60	8.59
Unrelated	Verb	1285.64	8.77

Table 2

By-subject means of verb production latency (in ms) by condition [within-subject standard errors] in Experiment 2.

Relatedness	DistractorType	mean	se
Related	Noun	1183.36	8.45
Related	Verb	1137.01	9.55
Unrelated	Noun	1138.16	5.94
Unrelated	Verb	1144.18	8.92

Table 3

By-subject means of subject noun production latency (in ms) by condition [within-subject standard errors] in Experiment 3.

Term	β	SE	|t|	p
Intercept	6.71	0.04	154	< 0***
Relatedness	0.03	0.01	2.21	< 0.05*
Trial order	-0.001	0.0002	4.80	0***

Table 4

Linear mixed effects model estimates of logged onset latency in Experiment 1. The random effects structure included the by-subject and by-item random intercepts, and by-subject and by-item random slopes for relatedness.

Term	β	SE	|t|	p
Intercept	7.14	0.02	270.87	< 0.00***
Relatedness	0.02	0.01	2.38	< 0.05*
Distractor category	0.01	0.01	0.63	0.53
Rel. x Dist. Cat	0.02	0.01	2.02	< 0.05*
Trial order	-0.00	0.00	21.20	0.00***

Table 5

Linear mixed effects model estimates of logged verb onset latency in Experiment 2. The random effects structure included the by-subject and by-item random intercepts, and by-subject and by-item random slopes for relatedness, distractor category, and the interaction between relatedness and distractor category.

Term	β	SE	|t|	p
Intercept	7.03	0.03	219.52	0.00***
Relatedness	0.02	0.01	1.69	0.11
Distractor category	0.02	0.01	2.08	< 0.05*
Rel. x Dist. Cat	0.03	0.02	2.04	< 0.05*
Trial order	-0.001	0.00007	13.35	0.00***

Table 6

Linear mixed effects model estimates of logged verb onset latency in Experiment 3. The random effects structure included the by-subject and by-item random intercepts, and by-subject and by-item random slopes for relatedness, distractor category, and the interaction between relatedness and distractor category.

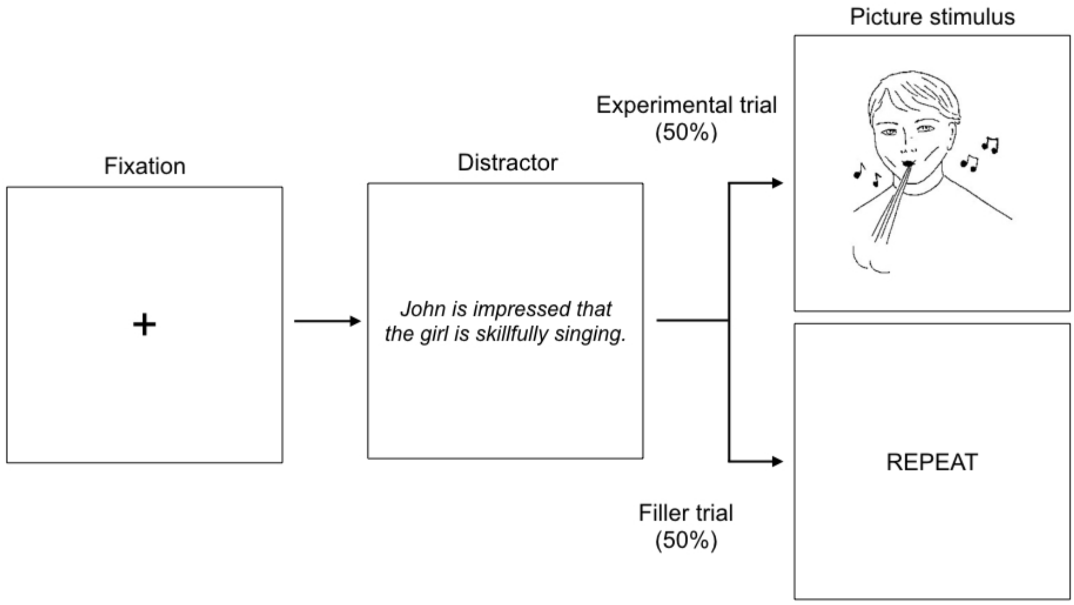

Figure 1. Graphical illustration of the basic task structure in the sentence-picture interference task.

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.	

Figure 2. Example distractor sentences used in Experiment 2. Note that all unrelated distractor sentences were also used as related distractor sentences in other trials of 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.

Figure 3. An example picture stimulus in Experiment 3.

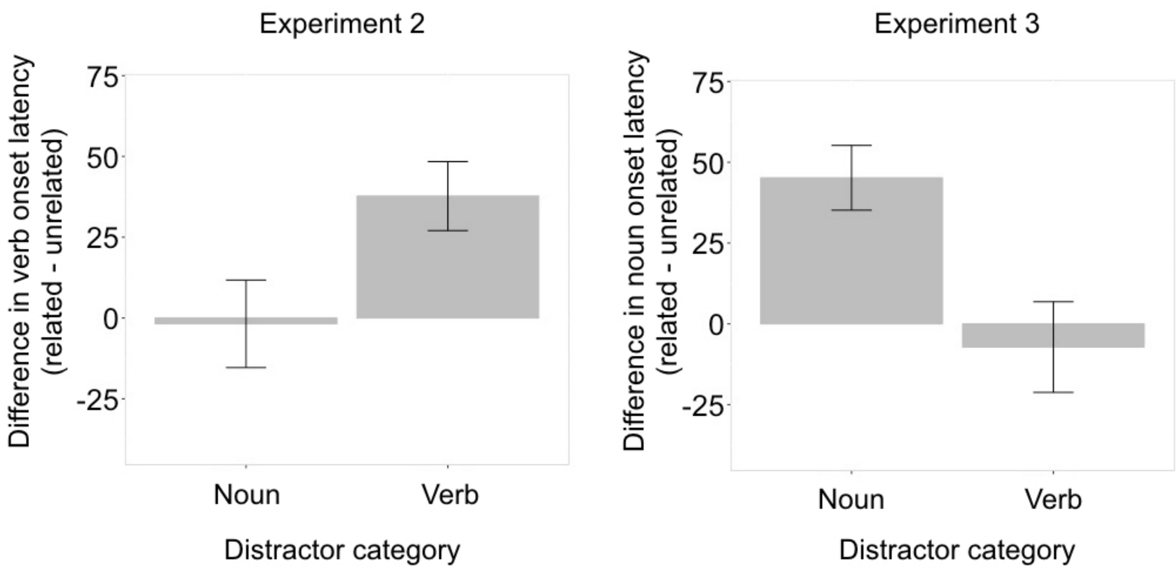

Figure 4. Interference effects in Experiment 2 and Experiment 3.