Cognitive Psychology 114 (2019) 101228

Contents lists available at ScienceDirect

Cognitive Psychology

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

Beyond linear order: The role of argument structure in speaking☆

Shota Momma[a,b,*], Victor S. Ferreira[a]

[a] University of California, San Diego, United States
[b] University of Massachusetts, Amherst, United States

ARTICLE INFO

Keywords:
Sentence production
Advance planning
Grammatical encoding
Argument structure
Extended Picture-word interference

ABSTRACT

The current study examines how speakers plan sentences in which two words that form hierarchical dependency relationships - arguments and verbs - appear far apart in linear distance, to investigate how linear and hierarchical aspects of sentences simultaneously shape sentence planning processes. The results of six extended picture-word interference experiments suggest that speakers retrieve sentence-final verbs before the articulation of their sentence-initial patient or theme arguments, but not agent arguments, and before retrieving sentence-medial nouns inside modifiers. These results suggest that the time-course of sentence planning reflects hierarchically-defined dependency relationships over and above linear structure.

1. Introduction

In sentence production, the time-course of sentence planning may reflect both linear and hierarchical structures of sentences. On the one hand, speakers may formulate sentences following surface word-order, to synchronize the retrieval and articulation of words as much as possible, thereby minimizing potential working memory cost (Christiansen & Chater, 2016; De Smedt, 1996; Iwasaki, 2010; Van & Dietrich, 2003; Slevc, 2011; Levelt, 1989; Ferreira & Dell, 2000). On the other hand, speakers may "look-ahead" to words that are arbitrarily distant, to encode hierarchically-defined dependency relationships (e.g., Momma, Slevc, & Phillips, 2016; Momma, Slevc, & Phillips, 2018). The current study examines how speakers plan sentences in which two word classes that form hierarchical dependency relationships - arguments and verbs - appear far apart in linear distance, to investigate how linear and hierarchical aspects of sentences simultaneously shape sentence planning processes.

1.1. The relationship between word-order and sentence planning

Sentences are specifically ordered sequences of words. Unlike hierarchical relationships, word-order is an observable property of linguistic input, so it is reasonable to assume that speakers learn to plan words according to the surface word-order of their language when developing their language skills. Planning words according to the surface word-order is also beneficial from the perspective of working memory economy. Speakers can avoid buffering words in memory if the order of planning is the same as surface word-order. In accordance with these learning and memory considerations, many models of sentence production assume that the time-course of lexical retrieval mirrors the surface word-order of sentences (Christiansen & Chater, 2016; De Smedt, 1996; Ferreira & Dell, 2000; Iwasaki, 2010; Kempen & Hoenkamp, 1987; Levelt, 1989; Slevc, 2011; Van & Dietrich, 2003; Dell, Oppenheim, & Kittredge, 2008).

☆ This research was supported in part by the National Institutes of Health grant R01-HD051030 to Victor Ferreira.

* Corresponding author at: Department of Linguistics, University of Massachusetts, Amherst, N408 Integrative Learning Center, 650 North Pleasant Street, University of Massachusetts, Amherst, MA 01003, United States.
E-mail address: smomma@ucsd.edu (S. Momma).

https://doi.org/10.1016/j.cogpsych.2019.101228
Received 28 November 2018; Received in revised form 23 July 2019; Accepted 25 July 2019
Available online 14 August 2019
0010-0285/ © 2019 Elsevier Inc. All rights reserved.

That is, speakers are often assumed to retrieve words sequentially, from 'beginning to end.' The models that adopt this assumption can be called sequential models, and wide varieties of sentence production models, including the models that invoke hierarchical representations that guide sentence planning, can be classified as sequential models. For example, Levelt (1989), based on the model proposed by Kempen and Hoenkamp (1987), describes a model of sentence production in which the order of lemma retrieval roughly follows the linear order of words as spoken. De Smedt (1996) formalized this property using the notion of information cascading across levels of processing - the idea that information at a higher level processing is sent to the next level of processing without delay. When applied across all processing levels - from message generation to articulation - the consequence of information cascading is the tight synchronization between planning and articulation, which, in ideal situations, results in sequential lexical retrieval.

Iwasaki (2010) adopted this idea and argued that even the lemmas of sentences in head-final languages could be planned sequentially. These models are generally consistent with the approach proposed by Ferreira and Dell (2000). They described a model of sentence production in which speakers produce sentences following what they called *the principle of immediate mention*. According to this view, speakers aim to say available words immediately. Dell et al. (2008), based on Gordon and Dell (2003), adopted a model of sentence production in which words are "activated incrementally, in sequence: if word *i* precedes word *j*, then word *i* is activated first (e.g., Dell, 1986; MacKay, 1982)". In all, the assumption that the order of word retrieval follows the surface word-order is widely held, implicitly or explicitly. Note, however, that sequential models are not necessarily *radically incremental*; radically incremental models do not assume that higher-level representations (like semantic or syntactic structures) guide sentence planning. We take radically incremental models to be empirically untenable (e.g., Christianson & Ferreira, 2005; Momma et al., 2016; Momma et al., 2018).

1.2. The relationship between grammatical dependencies and sentence planning

Linear order is certainly an important property of sentences, but one of the defining characteristics of human language is that sentences are hierarchically organized (Chomsky, 1986; Sag & Pollard, 1987; Kaplan & Bresnan, 1982). Hierarchical structures are essential in defining semantic and (morpho-) syntactic relationships between words in sentences. A prominent example involves semantic and syntactic dependencies between arguments and verbs. Arguments and verbs form dependency relationships, in the sense that the syntactic, semantic and morphological status of arguments often depends on verbs. For example, the relational meaning, grammatical function, morphological and syntactic case, syntactic and semantic categories, and even the presence or absence of arguments depends on the lexical properties of verbs. Consistent with these linguistic considerations, some prominent production models (e.g., Bock & Levelt, 1994; Ferreira, 2000) assume that verb lemmas must be selected to encode arguments grammatically. For example, if the only information that a speaker has access to is a thematic or semantic role information, it is unclear how they encode patient arguments as the subject, as in passive sentences and in sentences headed by unaccusative verbs. This claim is subsumed under the *verb-guidance hypothesis*, which claims that verbs are critically involved in determining the syntactic status of arguments. Critically for current purposes, the verb guidance hypothesis predicts that verbs should be selected before the encoding (and therefore the articulation) of their arguments, even when they appear last in a sentence. That is, under the verb guidance hypothesis, the time-course of sentence planning reflects argument-verb dependencies over and above surface word-order.

Strong versions of the verb-guidance hypothesis are challenged in the literature (Allum & Wheeldon, 2007; Iwasaki, 2010). Notably, Schriefers, Teruel, and Meinshausen (1998) conducted a series of experiments showing that verbs in German verb-final clauses do not need to be retrieved before speakers start to produce the subject of a sentence. Based on such evidence, it has been argued that verb lemmas are not necessary to encode the grammatical structures of sentences (Allum & Wheeldon, 2007; Iwasaki, 2010; Schriefers et al., 1998). Schriefers and colleagues' experiments were based on the *extended picture word interference* paradigm (Meyer, 1996; Momma et al., 2016; Momma et al., 2018; Momma, Slevc, & Phillips, 2015; Schriefers et al., 1998), in which speakers describe a picture in phrasal or sentential format, while ignoring a distractor that targets the non-initial words in an utterance. For example, speakers in Schriefers et al. (1998) produced a sentence like *the man fills the bucket* (in German), while ignoring a semantically related verb distractor like *empty* (in German). In their study, speakers produced such sentences in VS(O) or S(O)V word-order (where the presence of O depended on the transitivity of the verb). They found the verb interference effect on in verb-initial VS(O) sentences but not in verb-final S(O)V sentences in the onset latency measurement (see below for the detailed discussion of this method). Based on this absence of verb interference effect in the sentence onset latency, they suggested that speakers do not need to necessarily plan verbs in advance to encode sentences. This verb-independent encoding of arguments, sometimes referred to as the *conceptual guidance hypothesis*, serves to reduce speakers" memory cost. If verbs are not needed to encode their arguments, speakers need not hold the yet-to-be-spoken verbs in memory while retrieving intervening words until those verbs can be spoken.

However, there are theoretical and empirical reasons to think that verbs may still be critical in grammatical encoding, but selectively when encoding object arguments. In some linguistic theories, object arguments are considered to be the only true argument of verb roots (Kratzer, 1996; Kratzer, 2003). Under such a view, subject 'arguments' are actually an argument of an independent head that assigns agent roles (known as *little v* or *Voice*), so they are actually not an argument of verb at all. More generally, most linguistic theories assume contrasts between object and subject arguments with respect to their relationships with verbs. For example, object arguments and verbs in exclusion of subject arguments constitute verb phrases, but subject arguments and verbs in exclusion of object arguments do not (except in some grammatical theories like categorical grammar, Steedman, 2000). For example, in a sentence like *Mary likes the book*, *likes the book* is a constituent but *Mary likes* is not. Object arguments, but not subject arguments, are selected (i.e., subcategorized) by verbs (Chomsky, 1965; Haegeman, 1991), in the sense that the presence or absence of object arguments and the type of object arguments depends on the lexical properties of verbs. Selections restrictions of verbs that are irreducible to conceptual structures determine properties of object arguments (Grimshaw, 1990). For example, it is easy to find

verbs that are similar in meaning but differ in whether they require objects (e.g., *devour* requires an object argument but *eat* does not), but the same is not true for subject arguments. It is common that object arguments and verbs, in exclusion of subject arguments, form idioms (e.g., *kick the bucket*), but idioms that involve subject arguments and verbs in exclusion of object arguments are rare (Marantz, 1981), if they exist at all. Subject arguments exist independently of verbs in English and many other languages (there is even a formal principle encoding this contrast *Extended Projection Principle*, Chomsky & Keyser, 1982); verbs that do not assign meaning to subjects still require formal subjects (known as an expletive or pleonastic subject, e.g., *it rains*). Object arguments are easier to extract from embedded clauses in long-distance dependencies due to licensing from verbs (namely, *theta marking*, Chomsky, 1986). Generally speaking, it is widely assumed in linguistic theory that object arguments hold some special relationship to verbs that subject arguments do not. Consistent with this theoretical consideration, Momma et al. (2016) argued that Japanese speakers retrieve verbs before starting to speak object-initial sentences, but not before subject-initial sentences. Thus, both theoretical linguistic considerations and psycholinguistic evidence suggest that verbs may be critical for the grammatical encoding of their arguments, but only their object arguments.

More recent investigations have found that speakers retrieve verbs not just before the surface objects of sentences but also before underlying objects (semantic objects) of sentences. Using a similar method as Schriefers et al., Momma et al. (2015) showed that speakers plan verbs before speaking the subjects of passive sentences (which are semantically objects and hence are internal arguments) but not before the subject of an active sentence. Building on this finding, Momma et al. (2018) contrasted the timing of verb retrieval in two types of intransitive sentences: sentences headed by *unaccusative* verbs and sentences headed by *unergative* verbs. Unaccusative verbs are a type of intransitive verb whose sole argument is a theme or patient argument (e.g., *boil*, in that in *The octopus boils*, the octopus is not doing the boiling, but being boiled). Unergative verbs are intransitive verbs whose sole argument is an agent argument (e.g., *swim*, as in *The octopus swims*, the octopus is doing the swimming). Importantly, the subject argument of an unaccusative sentence, though appearing as the surface subject, exhibits semantic and syntactic properties typical of object arguments in many languages, including in English (see Perlmutter, 1968; Levin & Hovav, 1995; Alexiadou, Anagnostopoulou, & Everaert, 2004 among many others) and therefore is considered an internal argument in some linguistic theories (e.g., see Perlmutter, 1978). Momma and colleagues found that speakers showed the verb semantic interference effect in onset latencies in sentences headed by unaccusative verbs but not in sentences headed by unergative sentences. In contrast, they showed that speakers experienced verb semantic interference effects as they utter the pre-verbal words in sentences headed by unergative verbs but not in sentences headed by unaccusative verbs. Based on this temporal pattern of verb interference effects, they claimed that speakers retrieve verbs before speaking the subject of unaccusative verbs, but not before the subject of unergative verbs. Thus, the generalization about the timing of verb retrieval processes across the studies described above seems to be that speakers retrieve verbs before starting to speaking their *internal* arguments or their semantic objects - including stereotypical objects, passive subjects, and unaccusative subjects. Henceforth, we use the term *internal argument* for expository purposes, without theoretical commitments about whether the semantic object should also be treated as syntactic complement of the verb in the underlying representation of sentences.

1.3. Looking ahead to downstream verbs

The time-course of lexical retrieval processes can be guided by the linear structure of sentences, or by verb-argument dependencies. However, these two hypothesized factors can conflict with each other. Namely, arguments and verbs can appear (in principle) arbitrarily far apart in linear distance. For example, an indefinite number of words can intervene between the subject noun head and its unaccusative verb: *The octopus that John found in the grocery store next to the gas station…is boiling*. If speakers retrieve words sequentially, and if they need to retrieve verbs before speaking their internal arguments, it is predicted that speakers start speaking unaccusative sentences (and also passive sentences) only after they retrieved all the words in the sentence up to the verb (see the *head principle* by Martin & Freedman, 2001 for a related view). This predicts that speakers need to allocate processing time proportional to the number of words that intervene between an unaccusative verb's subject and that verb before they can start speaking an unaccusative subject. This prediction is intuitively implausible, and as far as we know has no empirical support.

There are at least two possible ways to avoid this potential need for extensive buffering of words. The first is to weaken the role of verbs during grammatical encoding. Advance verb retrieval may only occur when a verb appears close in linear distance to its internal arguments. For example, speakers may have some rough estimate of the complexity of an utterance (cf. Yamashita & Chang, 2001; Griffin, 2001), and retrieve verbs in advance only when the estimated complexity of the utterance is simple. When the utterance is estimated to be complex, they may start speaking without retrieving verbs in advance, assuming that the sentence can be continued coherently and grammatically. Under this view, speakers should not need to retrieve verbs to encode their internal arguments. This view requires sentence production models to have two different ways to encode internal arguments: verb-dependently (when the sentence is simple) and verb-independently (when the sentence is complex).

The alternative possibility is for sentence formulation mechanisms to retrieve verbs before speaking internal arguments, without retrieving words that occur in between, in order to robustly encode the dependency relationship between internal arguments and verbs while not overloading working memory with too many words. Under this view, preserving the correspondence between the order of mention and the order of lexical retrieval processes is not the prioritized goal of speakers. Speakers instead prioritize establishing the linguistic dependency between arguments and verbs, and they do so by systematically violating the correspondence between order of mention and order of planning. The advantage of this approach is the simplification of the models of grammatical encoding, in that internal arguments can always be encoded in a verb-dependent fashion, so there is no need to assume two different ways to encode internal arguments.

Figs. 1 and 2 illustrate the two mechanistic strategies for producing sentences with unaccusative verbs, assuming phrase-by-

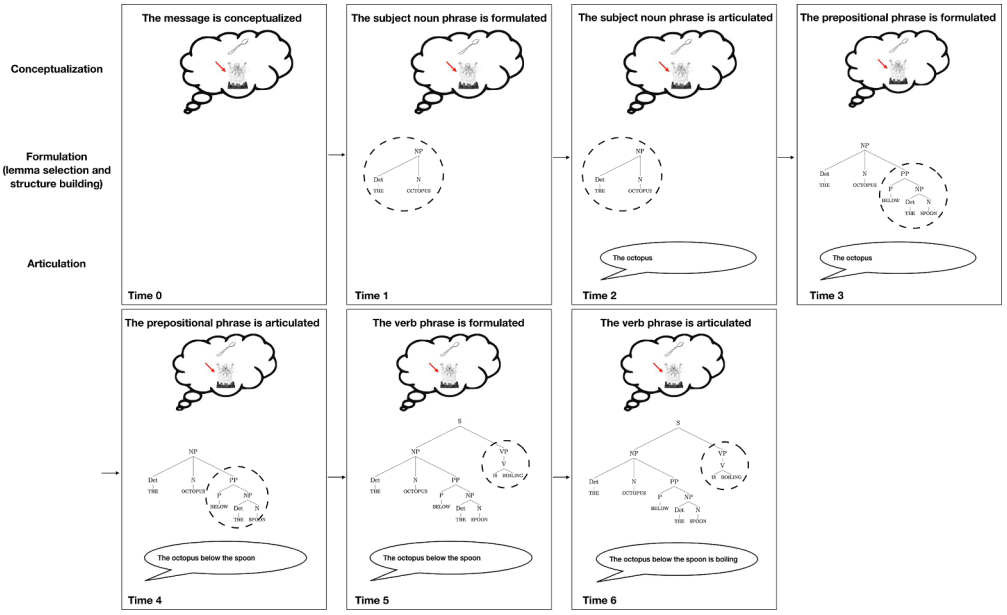

Fig. 1. Illustration of how a linear sentence planning mechanism generates the unaccusative sentence, *The octopus below the spoon is boiling*. Each panel represents a time step of the sentence planning process. The top row in each panel represents conceptual representations, which are assumed to be represented as a whole before sentence formulation begins. The middle row represents the syntactic representations. The bottom row represents the articulated part of the target sentence. The dotted circle represents the locus of speakers' attention. Note, in Time 2, formulation process commit to and articulate a subject NP without any certainty that the verb it depends on represents all NP properties (e.g., case) that have presumably been encoded into the NP.

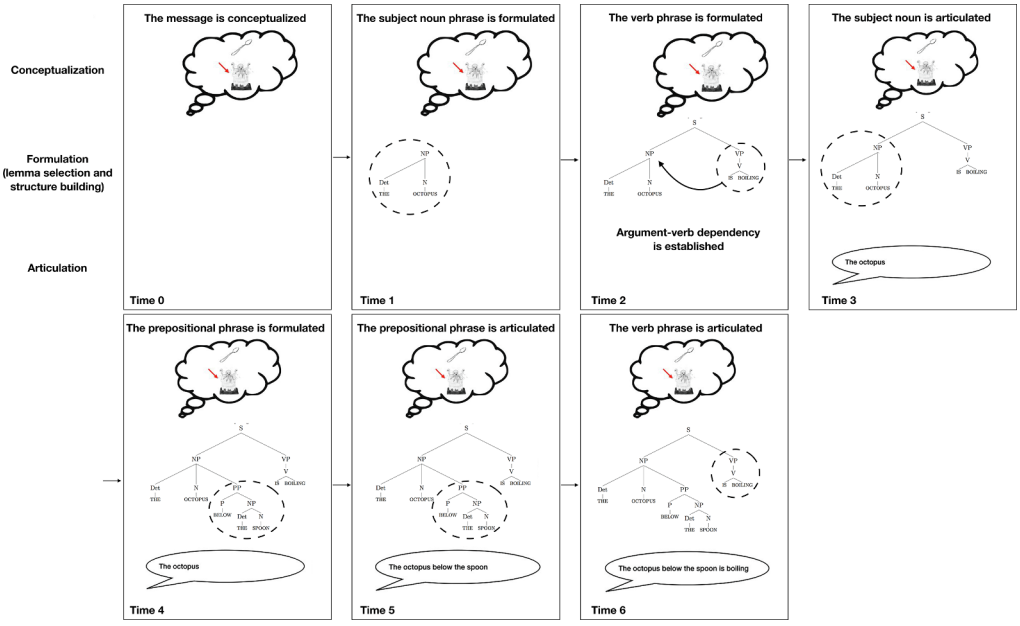

Fig. 2. Illustration of how a dependency-based sentence planning mechanism works for the unaccusative sentence, *The octopus below the spoon is boiling*. Each panel represents a time step of the sentence planning process. The top row in each panel represents conceptual representations, which are assumed to be represented as a whole before sentence formulation begins. The middle row represents the syntactic representations. The bottom row represents the articulated part of the target sentence. The dotted circle represents the locus of speakers' attention.

phrase planning and articulation of sentences (note that this assumption is adopted simply for illustrative purpose). Fig. 1 illustrates a strategy of the production mechanism that constructs the sentence linearly, without relying on verb guidance to encode internal arguments. Fig. 2, in contrast, illustrates the non-linear strategy of a production mechanism that prioritizes encoding the dependency between internal arguments and verbs. The current study examines which mechanistic strategy is adopted by speakers. To do so, we

build on the previous finding that speakers retrieve unaccusative verbs before starting to speak their internal arguments (Momma et al., 2018). We investigate the production of sentences in which internal arguments and unaccusative verbs are separated by another phrase that is not directly relevant to verb-argument dependency (e.g., *the octopus below the lemon is boiling*, see the description of the study design below). If lexical retrieval is indeed obligatorily sequential, speakers should not retrieve sentence-final unaccusative verbs without first retrieving all of the words in any intervening phrase. If, on the other hand, lexical retrieval is not obligatorily sequential, speakers may retrieve sentence-final unaccusative verbs in advance, without first retrieving all of the words in the intervening phrase.

1.4. Probing the time-course of lexical retrievals

Before we describe the details of the experimental design, we discuss how the time-course of lexical retrieval in sentence production can be studied experimentally. As briefly described above, the timing of retrieving a particular lemma in a sentence can be studied by characterizing the temporal profiles of *semantic interference effects* in what has been called the extended picture-word interference (ePWI) paradigm. ePWI is an extension of the widely used picture-word interference paradigm (Lupker, 1979). In the standard PWI paradigm, speakers produce a single word as a response to a picture stimulus, as they either hear or see a distractor word. Distractor presentation timing varies, but distractors are usually presented temporally close to the picture presentation onset. The classic finding using this paradigm is the *semantic interference effect*: speakers are slower to name pictures given distractors that are conceptually similar to the pictured referent (Lupker, 1979; Schriefers, Meyer, & Levelt, 1990; Roelofs, 1992; Vigliocco, Vinson, Lewis, & Garrett, 2004). This effect is considered to reflect the selection of lemmas, abstract linguistic representations that carry semantic and syntactic (but not phonological) information about the word (Kempen & Huijbers, 1983; Levelt, Roelofs, & Meyer, 1999). Interference does not seem to reflect conceptual preparation processes because an analogous interference effect is absent in picture categorization tasks (Schriefers et al., 1990). ePWI is an extended version of this paradigm, in which speakers produce not just a single word but a phrase or sentence (Meyer, 1996; Schriefers et al., 1998; Momma et al., 2016; Momma et al., 2018; Hwang & Kaiser, 2014; Schnur, 2011). Traditionally, ePWI has been used to test whether a particular word in a sentence is retrieved before the onset of the utterance that contains that word. However, this method can be further extended to measure the relative timing at which speakers retrieve particular words in sentences. Specifically, it is possible to measure the production duration of each word of a speakers' utterance and to measure the effect of distractor words on production duration. We use this word-by-word production duration measurement in the current experiments to characterize the temporal profile of semantic interference effects to assess the relative timing of sentence-medial noun planning versus sentence-final verb planning.

1.5. Current experiments

1.5.1. Hypotheses and predictions

The critical hypotheses are threefold. The first hypothesis is that speakers selectively retrieve sentence-final unaccusative verbs, but not sentence-final unergative verbs, before the onset of articulation of the subject nouns of those unaccusative verbs. This hypothesis is derived from the previous study by Momma et al. (2018), and also from linguistic considerations regarding the verb-argument relation discussed above.

The second and third hypotheses are that speakers retrieve (a) sentence-final unergative (but not unaccusative) verbs and (b) sentence-medial nouns inside subject-modifying adjunct phrases (henceforth adjunct nouns) on a just-in-time basis. These hypotheses are compatible with the widely accepted idea that planning and articulation can interleave within a single utterance.

These hypotheses generate three specific predictions in these ePWI experiments. The prediction based on the first hypothesis is that speakers should be slower to start speaking unaccusative sentences given a distractor word that is semantically related to unaccusative verbs. A distractor that is related to an unaccusative verb should slow the retrieval of that verb; if the subject of the unaccusative verb cannot be articulated until the unaccusative verb is retrieved, as described in Fig. 1, then subject retrieval should also be slowed, delaying the onset of sentence production. The prediction based on the second hypothesis is that speakers should not be slower to begin speaking unergative sentences given a distractor word that is semantically related to the unergative verb. Instead, speakers should elongate the production time of the words just preceding the unergative verbs. This is because subject noun phrases and unergative verbs do not hold the true argument-verb dependency, as discussed above, so the time course of sentence formulation is as described in Fig. 2. The prediction based on the third hypothesis is that speakers should not be slower to start speaking unaccusative or unergative sentences given a distractor word that is semantically related to an intervening adjunct noun. Instead, they should elongate the production time of the words immediately preceding the adjunct nouns. This is because, both in Figs. 1 and 2, the subject noun does not need any licensing from the subject modifying adjunct phrase to be articulated.

These three hypotheses together form a broader hypothesis that the time-course of sentence planning reflects verb-argument dependencies over and above the surface word-order of sentences.

1.5.2. Overview of the experiments

We report five experiments that involved different variants of ePWI tasks. In the first four experiments (Experiment 1, 2a, 2b, and 3), we used the ePWI paradigm to investigate when speakers retrieve sentence-final verbs and sentence-medial nouns (see the method sections below for more details), building on the previous study by Momma et al. (2018). Sentence-medial nouns were always inside a prepositional phrase headed by either *above* or *below*. Given pictures as in Fig. 3 and the instruction to describe the entity indicated by the red arrow, speakers consistently produced sentences like:

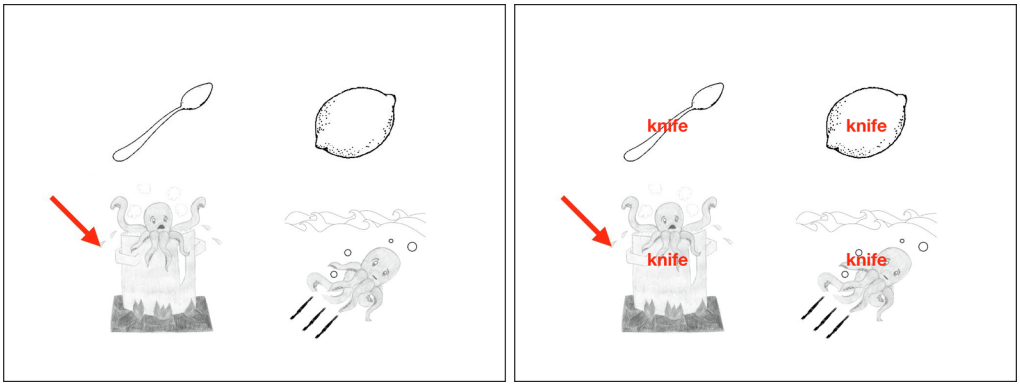

Fig. 3. Two example picture stimuli used in Experiment 1–4 with (right) and without (left) distractor words superimposed. Note that the redundant presentation of distractors on each component of the picture was to prevent speakers from visually ignoring the distractors. These pictures both elicited *The octopus below the spoon is boiling*. When the red arrow points to the octopus in the bottom-right corner, it elicited *The octopus below the lemon is swimming*. In half the stimulus lists, the position of two actions is switched. Across different picture stimuli, the action pictures corresponding to unaccusative verbs appeared both in the right and left the side of the screen roughly equally often.

(1) The octopus below the spoon is boiling [unaccusative]
(2) The octopus below the spoon is swimming [unergative]

The rationale for using the subject-modifying prepositional phrases are that (a) it is relatively easy to elicit using pictures and (b) that it introduces a noun that does not directly involved in the argument-verb dependency between the subject noun and their verb. Because the subject nouns do not depend on the sentence-medial nouns, these sentences are ideal for testing our hypothesis that sentence-final verb planning is prioritized over sentence-medial noun that the subject noun does not depend on. To measure when speakers retrieve the relevant lemmas, we presented visual distractor words that were conceptually related to either the noun head inside a subject modifying adjunct (e.g., *knife*) or to the verb (e.g., *melt*). The timing of the interference effect due to the related distractors (as compared to unrelated distractors) should correspond to when speakers retrieve the relevant lemmas. Experiments 1, 2a and 2b used essentially the same task structure, except that the timing of distractor presentation was different (150 ms before the picture onset in Experiment 1, at the same time as picture onset in Experiment 2a, 300 ms after the picture onset in Experiment 2b). Experiment 3 is a replication of the temporal pattern of verb interference effects in Experiment 1 and 2a, with two different timings of distractor presentation.

We conducted two additional experiments, where participants performed a standard ePWI task using the same set of event pictures (without object pictures) and distractor words as in Experiment 1–3. But in these last two experiments, speakers produced simple sentences, where the adjunct modifying the subject noun phrase was absent, like *The octopus is boiling/swimming*. This allowed Experiment 4 and 5 to serve as a test for the effectiveness of the distractor verbs used in Experiment 1–3. Experiment 4 and 5 were the same in design, except that Experiment 5 involved a stop-signal task (see below), to make the experimental task more comparable to previous published experiment (Momma et al., 2018).

In all experiments, we used two measures to examine the timing of noun and verb retrieval. First, we focused on the subject noun onset latency rather than utterance onset latency because speakers use different forms of *the* (e.g., *the* vs. *thee*) in order to signal a suspension of speech (Fox & Clark, 1997), so that the simple speech onset latency may not be a suitable measure of the processing cost associated with advance sentence planning. Therefore, the measurement that is likely to be more suitable for estimating the processing cost associated with advance sentence planning is the combined measure of speech onset latency and the production time of the initial determiner. We henceforth call this first region of interest *before-subject noun* region.

Second, we defined the second region of interest to be the total production time of the two words prior to the critical word (e.g., for *the octopus above the spoon is boiling* in Experiment 1–3, *spoon is* in the verb distractor conditions, and *above the* in the noun distractor condition, and for *the octopus is swimming* in Experiments 4 and 5 *octopus is* in the verb distractor conditions). This decision was based on the hypothesis that speakers retrieve unergative verbs and adjunct nouns on a just-in-time basis. Previous studies suggest that it takes roughly 250–450 ms from the initiation of the lemma selection process (the process of interest) to initiate the articulation in picture naming studies investigating single-word production processes (Indefrey & Levelt, 2004). Because it takes roughly 100–250 ms on average to produce a single syllable, the average temporal locus of the semantic interference effect can be estimated to be roughly 2–3 syllables before the production onset of the word in question, if speakers retrieve words on a just-in-time basis. In the current experiments, 2–3 syllables before the critical words correspond to the adjunct noun head for verbs and preposition for adjunct nouns. However, this estimate assumes fluent speech with a uniform speech rate. In order to accommodate minor disfluencies, changes in speech rate, and some random variability in when words are retrieved, the total production time of the two words preceding the critical word were used to assess whether speakers retrieve unergative verbs and adjunct nouns on just-in-time basis. We henceforth call this second region of interest *before-verb* region in the verb distractor conditions, and *before-adjunct-noun* region in the noun distractor conditions.

2. Experiment 1

The goal of Experiment 1 was to test whether speakers selectively retrieve unaccusative verbs (e.g., *boil*) before the onset of sentence-initial subject nouns and whether speakers retrieve adjunct nouns (e.g., *spoon*) and unergative verbs (e.g., *swim*) later in their utterances, on a *just-in-time* basis. Based on the hypothesis that speakers retrieve unaccusative verbs before the production of their subject arguments, but they retrieve unergative verbs and adjunct nouns on a just-in-time basis, three predictions were generated. First, the unaccusative verb interference effect, but not the unergative verb interference effect, should be observed before the onset of the subject noun. Second, the unergative verb interference effect should be observed in the regions just preceding the unergative verbs (Momma et al., 2018). Third, the adjunct noun interference effect should be observed in the regions neighboring the noun but not in the sentence onset.

2.1. Method

2.1.1. Participants

Sixty undergraduate students at the University of California, San Diego participated in the experiment for course credit. Two participants were replaced because their first language was not English, according to self-report. Another participant was also replaced because of low accuracy (< 50% in at least one of the conditions). All remaining participants reported that they learned English as their first language. Informed consent was obtained from each participant before the experimental session.

2.1.2. Materials

Twenty-four event pictures with a person or animal and twenty-four object pictures were combined to yield forty-eight pictures like Fig. 3. Twelve of the event pictures corresponded to unaccusative sentences (e.g., *the octopus below the spoon is boiling*). The remaining twelve event pictures corresponded to unergative sentences (e.g., *the octopus below the lemon is swimming*). The unaccusative and unergative verbs were mostly chosen based on transitivity alternation tests and were confirmed using additional tests used in Momma et al. (2018). The average log frequency (per million) was well matched between unaccusative and unergative verbs (unaccusative: M = 8.90, SD = 1.03; unergative: M = 8.61, SD = 1.68; $t(22) = 0.27$, $p = .61$). Each of the twenty-four complex pictures contained two events (e.g., *boiling* and *swimming* events) sharing the same event participant (e.g., *octopus*), and two object pictures (e.g., *lemon* and *spoon*) all taken from the UCSD International Picture Naming (IPNP: Szekely et al., 2004) database. In half of the pictures, the event pictures were placed on the top half of the display. In another half of the pictures, event pictures were placed on the bottom half of the display.

For each picture, the positioning of the two object entities (e.g, *spoon* and *lemon*) was switched to yield another twenty-four complex pictures. These two versions of the complex pictures were distributed across two different experimental lists, so the sets of words preceding the critical verbs (e.g., *the octopus below the spoon/lemon*) were identical (across subjects) between unaccusative and unergative conditions. For each version of the pictures, a red arrow pointed to one of the action pictures, so participants said either unergative or unaccusative sentences depending on the action that the red arrow pointed to. Furthermore, based on these two experimental picture sets, we created two versions of the lists with different random orders of trials. This yielded four different stimulus lists. Finally, based on these four different stimulus lists, we created an additional four stimulus lists by reversing the order of trials of each list. Thus, we used a total of eight different stimulus lists, and participants were distributed roughly evenly across these lists. The entire set of target sentences, along with related and unrelated distractor words, is available in the Appendix. All the pictures used here can be downloaded from https://osf.io/vp2kf/?view_only=b2760a4d347449ffad30936a9f67f53f.

For each picture, the related distractor words were first chosen based on intuition, and then relatedness was verified using a cosine similarity measure from Latent Semantic Analysis (Landauer & Dumais, 1997). Although this measure is far from perfect in predicting the semantic interference effect, the resources for computing this are easily accessible (e.g., http://lsa.colorado.edu/) and has been frequently used in assessing semantic similarity, including in the norming of previous picture-word interference study assessing the current issues (e.g., Momma et al., 2018). For verbs, we specifically chose distractors in such a way that (a) the average relatedness between the subject nouns and the verb distractors did not differ between unaccusative and unergative conditions, (b) the average relatedness between the distractors and the target verbs in unaccusative and unergative conditions are approximately equal. The mean cosine similarity for related verb distractors is.28 ($SE = 0.04$) in the unaccusative conditions and.32 ($SE = 0.05$) in the unergative conditions. Thus, the cosine similarity measures were well matched between unaccusative and unergative conditions ($t(22) = 0.14$, $p > .5$). The related verb distractors were re-paired with other pictures within the same verb type to yield unrelated distractors. This procedure ensures that the set of related and unrelated distractor words are identical, so no first-order properties of distractor words (e.g., frequency, word length, imageability, etc.) can explain any potential differences between the related and unrelated conditions. The mean cosine similarity between the target and the unrelated distractors was comparable between unergative verbs ($M = .10$, $SE = 0.01$) and unaccusative verbs ($M = .09$, $SE = 0.01$, $t(22) = 0.23$, $p > .5$). The cosine similarity difference between related and unrelated verb distractors was statistically reliable in the unaccusative conditions ($t(11) = 4.23$, $p < .001$) and in the unergative conditions ($t(11) = 4.21$, $p < .001$). For noun distractors, we also chose related distractors based on intuition and then verified the relatedness judgment using LSA. The target-distractor noun pairs were identical between unaccusative and unergative conditions. For related noun distractors, the mean cosine distance between the target noun and the distractor nouns was.37 ($SE = 0.04$). Again, these related distractors were re-paired to create unrelated target-distractor pairs. For unrelated noun distractors, the mean cosine distance between the target and distractor nouns was.08 ($SE = 0.02$). The cosine similarity difference between related and unrelated noun distractors was statistically reliable ($t(23) = 6.65$, $p < .0001$).

2.1.3. Design

In Experiment 1, we manipulated three independent variables. We manipulated the type of verbs used in target utterances (VerbType: unaccusative vs. unergative), the type of the distractor words (DistractorType: Noun vs. Verb), and the relatedness of distractor words to the target (Relatedness: Related vs. Unrelated). Thus, the experiment adopted a 2 (VerbType) × 2 (DistractorType) × 2 (Relatedness) within-subject design. As discussed above, speakers uttered sentences that contained either an unergative verb (e.g., *the octopus below the spoon is swimming*) or an unaccusative verb (e.g., *the octopus below the spoon is boiling*) while seeing a noun distractor that was related (e.g., *knife*) or unrelated (e.g., *apple*), or verb distractor that was related (e.g., *melt/run*) or unrelated (e.g., *fall/smile*). No distractor-target pair started with the same syllable or rhymed. Following Momma et al. (2018), we used a stimulus onset asynchrony (SOA) of −150 ms. That is, the distractor words appeared 150 ms before the onset of the picture presentation.

2.1.4. Procedure

First, participants studied event pictures with the corresponding event and object descriptions using a picture booklet containing all experimental picture components (e.g., just one quadrant of the full scene shown in Fig. 3). Each picture was presented with the written target sentence, and participants were instructed to study the picture until they felt comfortable describing each picture using the target sentence. This familiarization process was to ensure that speakers say the sentences needed in the experimental session and is similar to a common practice in single-word picture-word interference studies (e.g., Schriefers et al., 1990), and it may even be a necessary procedure for obtaining reliable semantic interference effects (Collina, Tabossi, & De Simone, 2013). Following this familiarization session, participants practiced describing the complex pictures that they saw in the following experimental session. In this practice session, participants were instructed to first find a red arrow in the picture, and describe which of the two participants (indicated by the red arrow) is doing what action. After one practice for each picture, participants proceeded to the experimental session.

Each experimental trial started with a fixation cross for 500 ms. After the fixation cross, a full scene as in Fig. 3 was presented with a short click sound that was later used to identify the onset of the picture in the audio recording. In this experiment, distractor appeared on the screen 150 ms before the picture is presented. The picture remained on the screen for 5000 ms, and the distractor words remained on the screen for 2000 ms. The entire experiment session was audio recorded, and research assistants then transcribed the audio recording.

2.1.5. Analysis

The transcriptions and audio files corresponding to each individual trial were aligned using an automatic text-to-speech forced alignment algorithm (P2FA, Yuan & Liberman, 2008). From the output of the text-to-speech alignment algorithm, the speech onset latency and the production time of each word (duration and any potential pause before the onset of the next word) were extracted.

Based on these data, we computed two measurements that reflect the production time of two regions of interest: the onset latency of subject noun head (i.e., the onset latency of the sentence plus the duration of the sentence-initial determiner) for the before-subject noun region, and the total production time of the two words that precede each critical word (i.e., the total production time of *above/below the* in the noun distractor conditions and, e.g., *spoon is* in the verb distractor condition) for the before-adjunct noun region or the before-verb region.

For each region of interest, using *R* and *glmer* function of *lme4* (Bates, Mächler, Bolker, & Walker, 2015), we conducted a generalized mixed effects model to fit the production time using the Inverse-Gaussian distribution (also known as Wald distribution), using the identity link function (Lo & Andrews, 2015), with Relatedness, VerbType, and the Relatedness × VerbType interaction as fixed effects. Any trials containing overt hesitations (e.g., *um*) and errors (i.e., any utterances deviating from the target sentences) were not included in the analysis. All categorical variables (Relatedness, VerbType, and Region) were centered (i.e., −0.5 vs. 0.5). The random effect structures in all models were maximal (Barr, Levy, Scheepers, & Tily, 2013). In case of convergence failure, the random effects structures were simplified by successively removing the by-participant or by-item slopes that accounted for the least amount of variance, until model convergence. The rationale for conducting Inverse Gaussian analyses is that production time (both onset latency and production duration) are highly skewed, and there is an *a priori* reason to believe that verb semantic interference effects in sentence production resides in slow trials (Momma et al., 2016). Transforming production times and fitting them assuming a Gaussian distribution may obscure the true underlying pattern, because log-transformation of production time is likely to disproportionately 'shrink' effects that reside in slow trials. Thus, Gaussian mixed effects analysis based on the transformed production time is likely to be inadequate for reliably detecting potential semantic interference effects.

Finally, in case a significant interference effect was found, we conducted an additional analysis that aimed to characterize the time-course of the interference effects. In this time-course analysis, we compared the magnitude of an interference effect in the region that showed the interference effect in question with the other region that did not show the interference effect (e.g., if an adjunct noun interference effect is found in the before-adjunct region, the magnitude of this effect was compared to the magnitude of the interference effect in the before-subject noun region). To account for the scale difference, we first z-transformed the production time by each participant by each region, subsetted the data to the relevant conditions, and then constructed the linear mixed effects model on the z-scored production time. Because z-scores can be negative, we could not meaningfully use the inverse-Gaussian analysis, which cannot accommodate the negative scores. Thus, we used a typical linear mixed effect models using lmer. This model included Region (the before-subject noun region vs. before-adjunct noun region), Relatedness, VerbType and interaction between these factors as fixed effects. The construction of the random effects structure followed the same procedures as above.

Table 1
The by-subject mean raw production time for each word in Experiment 1, along with the error rate by each condition.

DistractorType	VerbType	Relatedness	Onset	The	octopus	below	the	spoon	is	accuracy %
Verb	Unacc	Related	1257	187	519	407	148	547	273	88
Verb	Unacc	Unrelated	1206	168	526	410	157	547	293	88
Verb	Unerg	Related	1216	166	509	405	158	532	246	87
Verb	Unerg	Unrelated	1218	190	508	399	161	537	247	91
Noun	Unacc	Related	1247	205	518	436	185	562	297	85
Noun	Unacc	Unrelated	1242	178	510	411	152	561	261	88
Noun	Unerg	Related	1210	179	499	422	187	565	267	88
Noun	Unerg	Unrelated	1217	177	495	402	157	534	233	90

2.2. Results

11.96% of all data points (689 of 5760 trials) came from erroneous trials (i.e., deviated from target sentences), and so were excluded from the subsequent timing analyses. The word-by-word mean production times and accuracy rates by each condition are summarized in Table 1. Fig. 4 is a difference plot showing the semantic interference effect for each word by distractor type.

2.2.1. The effect of verb distractors on production time

The production times in the two regions of interest by each condition are shown in Fig. 5. When speaking unaccusative sentences with verb distractors, speakers were 72 ms slower to start speaking the subject head noun in the related distractor condition than in the unrelated distractor condition. In comparison, when speaking unergative sentences with verb distractors, speakers were, if anything, 26 ms *faster* to start speaking the subject noun head in the related distractor condition than in the unrelated distractor condition. As can be seen in Table 2, the analysis of the subject noun onset latency revealed a two-way interaction between Relatedness and VerbType ($\hat{\beta} = 124$, $SE = 39$, $z = 3.16$, $p = .001$). There was no main effect of Relatedness or VerbType ($ps > .57$). To understand the nature of this interaction effect, we conducted pairwise comparisons by VerbType. This analysis revealed that the effect of Relatedness was significant in the unaccusative conditions ($\hat{\beta} = -73$, $SE = 34$, $z = -2.16$, $p = .03$), but not in the unergative conditions ($\hat{\beta} = 52$, $SE = 33$, $z = 1.58$, $p = .11$).

However, Experiment 1 provides no evidence as to when speakers retrieve unergative verbs. As can be seen in Table 2 and Fig. 5, the main effect of Relatedness and the interaction between Relatedness and VerbType were not significant in the before-verb region (all $ps > .33$). The main effect of VerbType was marginally significant ($\hat{\beta} = -58$, $SE = 30$, $z = -1.94$, $p = .053$). This marginally significant effect of VerbType is not of our main interest, but it might suggest that speakers spent more time producing the pre-verbal words in the unaccusative conditions than in the unergative conditions. Thus, we found no evidence that speakers were affected by related verb distractors at any point in the unergative conditions.

Given the robust unaccusative verb interference effect in the before- subject noun region, we conducted an additional analysis that aimed to characterize the time-course of the unaccusative interference effect. This analysis revealed that the two-way interaction

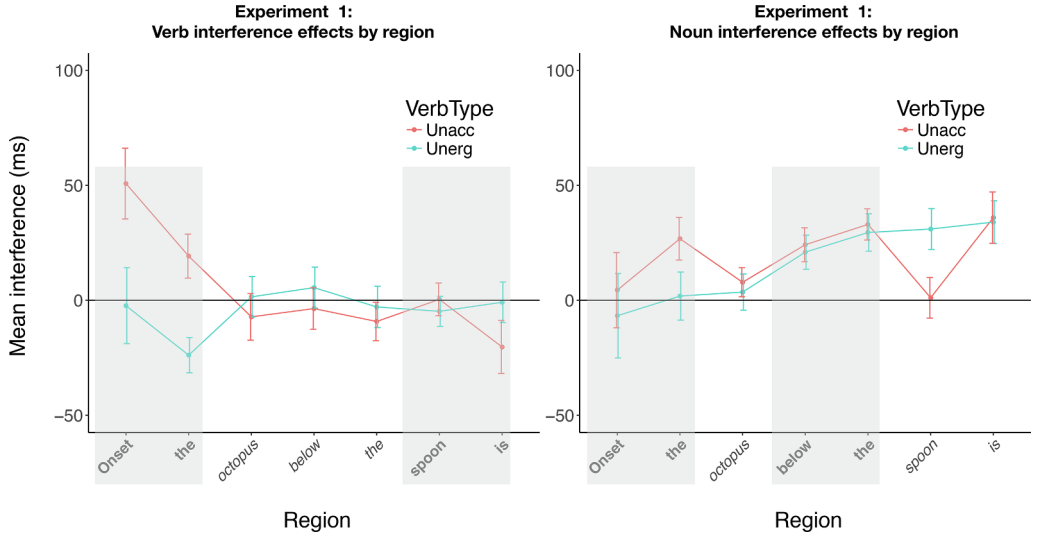

Fig. 4. Interference effects (in raw production time) on region-by-region production time by VerbType in the verb distractor condition (left) and the noun distractor condition (right) in Experiment 1. Error bars represent standard error of the mean. Gray areas represent the regions of interest for each type of distractor types.

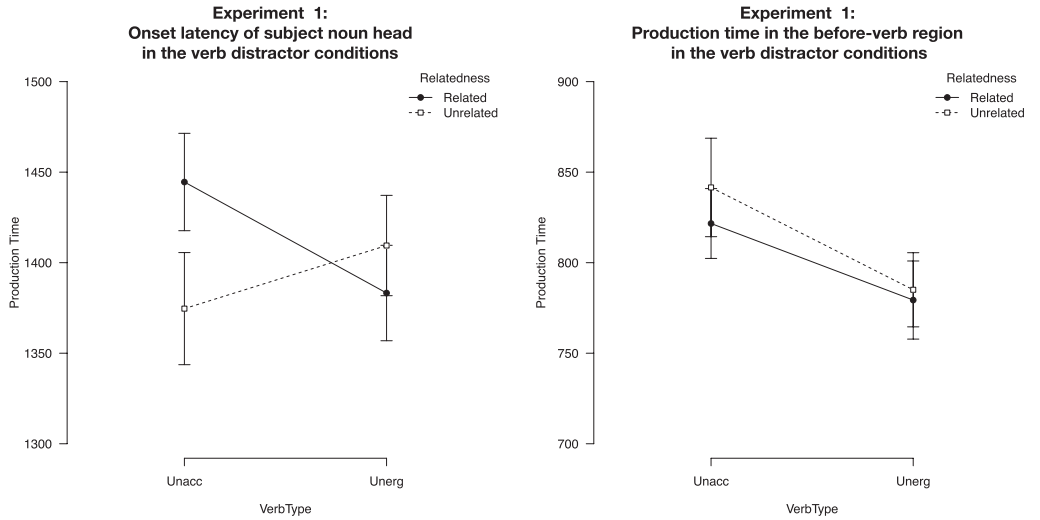

Fig. 5. Production time in two regions of interest in the verb distractor conditions in Experiment 1. Error bars represent 95% within-subject confidence intervals.

Table 2
Results of mixed effects model analyses on the two regions of interest in the verb distractor conditions in Experiment 1.

	Estimate	SE	z	p
Before-subject noun region				
Intercept	1607	49	32.67	< .001***
Relatedness	−10	27	−0.39	.695
VerbType	−22	40	−0.56	.579
Relatedness × VerbType	124	39	3.16	. 001**
Before-verb region				
Intercept	879	25	35.49	< .001***
Relatedness	3	16	0.16	.871
VerbType	−58	30	−1.94	.053
Relatedness × VerbType	−25	39	−0.65	.514

between Relatedness and Region was significant ($\hat{\beta} = -0.19$, $SE = 0.08$, $t = -2.46$, $p = .01$), suggesting that the interference effect was stronger in the before-subject noun region than in the before-verb region. No other effects were significant ($ps > .15$).

2.2.2. The effect of noun distractors on production time

The analysis of the noun distractor conditions in subject noun onset latency yields no evidence that speakers (mostly) retrieved the adjunct noun head before starting to speak the subject head noun. The production times in the pre-critical region of interest in each condition are shown in Fig. 6. Speakers were not reliably faster or slower to start speaking the subject head noun with the related distractor nouns. As can be seen in Table 3, the main effect of Relatedness, and the interaction between Relatedness and VerbType were not significant (all $ps > .24$). The main effect of VerbType was significant ($\hat{\beta} = -50$, $SE = 23$, $z = -2.14$, $p = .03$), suggesting that speakers were slower to start speaking the subject noun in the unaccusative than in the unergative conditions. In contrast, speakers were about 54 ms slower to articulate the before-adjunct noun region with related distractor nouns. As can be seen in Table 3, the main effect of Relatedness in the before-adjunct noun region was highly significant ($\hat{\beta} = -55$, $SE = 11$, $z = -5.21$, $p < .001$). There was no evidence that this effect differs by VerbType, as the interaction between Relatedness and VerbType was not significant ($p > .68$). The effect of VerbType was not significant either ($p > .75$). Thus, under the assumption that the semantic interference effect reflects the lemma retrieval process, speakers seem to engage in the lemma selection of adjunct head noun on a just-in-time basis.

Given the robust adjunct noun interference effect in the before-adjunct noun region, we conducted an additional analysis that aimed to characterize the time-course of the interference effect (see Analysis section above for details). This analysis revealed that the two-way interaction between Relatedness and Region was significant ($\hat{\beta} = 0.20$, $SE = 0.05$, $t = 3.58$, $p < .001$), suggesting that the interference effect was stronger in the before-adjunct noun region than in the before-subject noun region. The main effect of Relatedness was also significant ($\hat{\beta} = -0.14$, $SE = 0.04$, $t = -5.00$, $p < .001$), but we avoid interpreting this main effect due to the presence of the higher-order interaction. No other effects were significant ($ps > .06$).

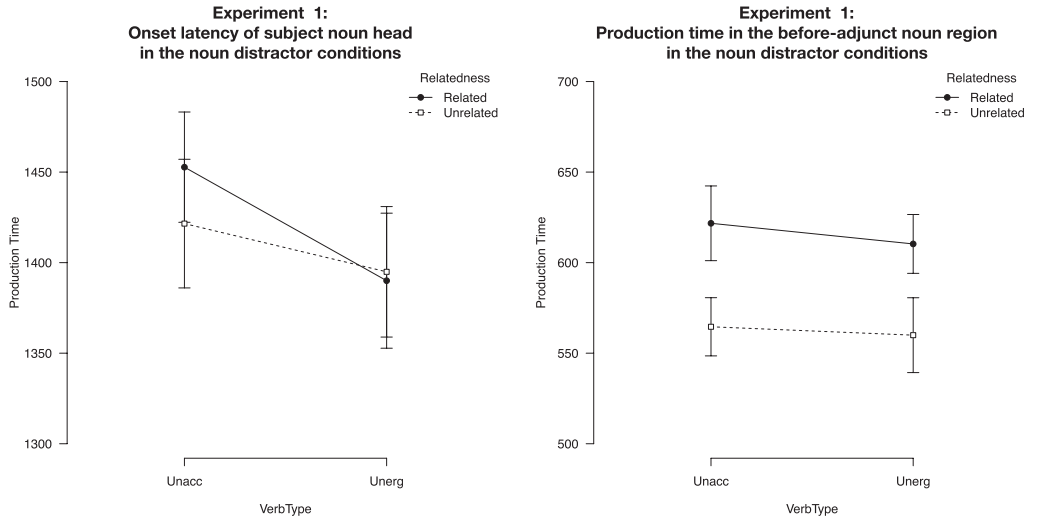

Fig. 6. Production time in two regions of interest in the noun distractor conditions in Experiment 1. Error bars represent 95% within-subject confidence intervals.

Table 3
Result of mixed effects model analyses on the two regions of interest in the noun distractor conditions in Experiment 1.

	Estimate	SE	z	p
Before-subject noun region				
Intercept	1588	24	66.29	< .001***
Relatedness	−20	19	−1.05	.293
VerbType	−50	23	−2.14	.033
Relatedness × VerbType	16	22	0.71	.475
Before-adjunct noun region				
Intercept	647	23	28.49	< .001***
Relatedness	−55	11	−5.21	< .001***
VerbType	−8	26	−0.32	.752
Relatedness × VerbType	5	14	0.40	.688

2.3. Discussion

Experiment 1 yielded two main findings. First, the semantic interference effect on unaccusative verbs, but not on unergative verbs, was found in the onset latency of the subject head noun, the measure that reflects advance planning. Second, the semantic interference effect on adjunct nouns was not found in the onset latency of the subject head noun, but found later, in the before-adjunct noun region. Note that this finding contrasts with previous claims that speakers retrieve the nouns of the first phrase of a sentence (Smith & Wheeldon, 1999; Martin & Freedman, 2001), and is more consistent with the claim that speakers need not retrieve a second noun inside the first phrase (Griffin, 2001) until just before articulation.

It is important to note that the absence of evidence for the adjunct noun interference in the onset latency does not constitute evidence that speakers do plan the adjunct nouns in advance in any trials. Thus, we do not claim that speakers always retrieve adjunct noun on a just-in-time basis. Instead, we argue that speakers retrieve adjunct nouns on a just-in-time basis more consistently than before the sentence onset (based on the two-way interaction between Relatedness and Region in the noun distractor condition). In all, Experiment 1 suggests that speakers can retrieve sentence-final unaccusative verbs without necessarily retrieving the sentence-medial adjunct noun. This order of lexical retrieval is directly in line with the hypothesis that the time-course of sentence planning reflects verb-argument dependency relationships at the expense of the linear relationship between words.

However, in Experiment 1, there was no evidence of semantic interference effects on unergative verbs anywhere in the utterance. There are at least three possible reasons for this null result. First, it is possible that our choice of distractor verbs for unergative verbs was not adequate. Although we used latent semantic analysis (Landauer & Dumais, 1997) to match the semantic relatedness of distractors in the unaccusative and the unergative conditions, the cosine-distance obtained from the latent semantic analysis measure is by no means a perfect predictor of the magnitude of semantic interference. Second, it is also possible that the activation of the unergative distractor dissipated before speakers needed to engage in unergative verb retrieval processes. If the activation of distractors dissipates over time (Bloem, van den Boogaard, & La Heij, 2004), it is expected that speakers might show no semantic interference effect on words that they retrieve late in their utterances. Third, it is possible that individual speakers are variable in when they retrieve unergative verbs, so the unergative verb interference effect does not appear in any single region. The subsequent

experiments evaluate these possibilities.

3. Experiments 2a and 2b

The primary goal of Experiment 2 was to examine why there was no unergative interference effect in Experiment 1. One potential reason was the long interval between the distractor presentation and the hypothesized timing of unergative verb planning. Thus, in Experiments 2a and 2b, we delayed the timing of distractor presentation relative to Experiment 1 by 150 ms (in 2a) and by 450 ms (in 2b). That is, Experiment 2a used an SOA of 0 ms, and Experiment 2b used an SOA of 300 ms. Additionally, Experiment 2a and 2b aimed to replicate the adjunct noun interference effect observed in the before-adjunct noun region in Experiment 1.

Note that the unaccusative verb interference effect was not necessarily predicted in Experiment 2a and 2b. The reason is that SOA is known to be a critical factor that modulates semantic interference effects in single-word picture naming studies (Schriefers et al., 1990). Schriefers and colleagues have shown that the semantic interference effect in a single-word picture interference task is most pronounced with an SOA of −150 ms (as in our Experiment 1). If unaccusative verbs are planned before sentence onset, as hypothesized here, it is critical that distractors are presented early enough to elicit semantic interference effects. Thus, given the delay in the distractor presentation in Experiment 2a and 2b, the unaccusative semantic interference effect may not be present.

3.1. Methods

3.1.1. Participants

Ninety-nine (forty-eight in Experiment 2a, fifty-one in Experiment 2b) undergraduate students at the University of California, San Diego participated in the experiment for course credits. Two participants in Experiment 2a and four participants in Experiment 2b were replaced due to low accuracy (< 50% in any condition). All participants reported that they learned English as their first language. None participated in Experiment 1. Informed consent was obtained for each participant before the experimental session.

3.1.2. Materials, procedure and analysis

The same materials, procedures, and analyses were used as in Experiment 1, except that the timing of distractor word presentation was delayed by 150 ms (to be a 0 ms SOA) in Experiment 2a, and 450 ms (to be a 300 ms SOA) in Experiment 2b.

3.2. Results of Experiment 2a (SOA = 0 ms)

13.78% of all data points (635 out of 4608 trials) were excluded from the subsequent analysis as having come from erroneous trials. The word-by-word mean production time, along with the accuracy rate (in %), by each condition is summarized in Table 4. Fig. 7 is the difference plot visualizing the semantic interference effect for each word.

3.2.1. The effect of verb distractors on production time

First, we report the analysis of the verb distractor conditions. The production times in the two regions of interest by each condition are shown in Fig. 8. As can be seen in Table 5, this analysis provides little evidence that speakers experienced semantic interference effects before they started to speak the subject noun. The main effect of Relatedness was marginally significant ($\hat{\beta} = -44$, $SE = 23$, $z = -1.92$, $p = .06$). Given this marginally significant effect, and given that we were interested in the difference between unaccusative and unergative conditions, we conducted the pairwise comparisons testing the effect of Relatedness by each VerbType. However, this analysis revealed the marginally significant effect of Relatedness in the unaccusative condition ($\hat{\beta} = -47$, $SE = 27$, $z = -1.77$, $p = .08$) and insignificant effect in the unergative condition ($\hat{\beta} = -42$, $SE = 27$, $z = -1.53$, $p = .12$). There was no significant interaction between Relatedness or VerbType ($p = .85$). There was also no evidence that the onset latency of the subject noun differed by VerbType: the main effect of verb type was not significant ($p = .87$). Thus, unlike in Experiment 1, we have only weak evidence that speakers experienced semantic interference before the onset of the subject noun. However, this is not necessarily in conflict with the results of Experiment 1, because the timing of distractor presentation was different and such a difference is known to influence whether speakers show semantic interference effects in single-word naming tasks (Schriefers et al., 1990).

Table 4
The by-subject mean production time in Experiment 2a and the accuracy (in %) by each condition.

DistractorType	VerbType	Relatedness	Onset	The	octopus	below	the	spoon	is	accuracy %
Verb	Unacc	Related	1246	210	505	412	169	531	283	86
Verb	Unacc	Unrelated	1212	200	508	418	176	528	291	86
Verb	Unerg	Related	1243	200	491	391	164	524	260	87
Verb	Unerg	Unrelated	1206	205	480	391	152	516	260	90
Noun	Unacc	Related	1265	236	498	436	225	558	297	82
Noun	Unacc	Unrelated	1246	235	517	408	173	521	272	88
Noun	Unerg	Related	1227	218	504	424	198	548	269	87
Noun	Unerg	Unrelated	1220	209	481	408	179	522	245	90

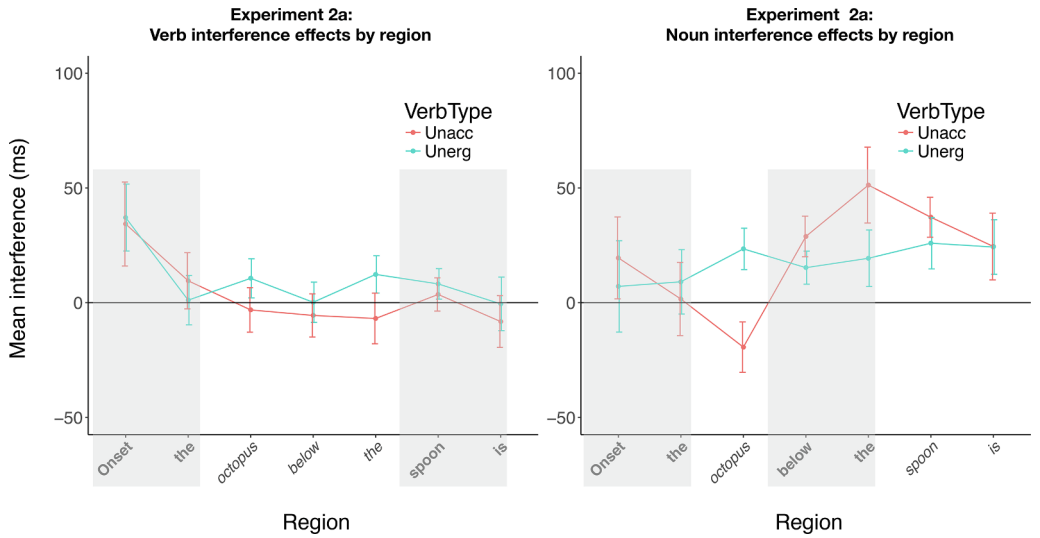

Fig. 7. Interference effect (in raw production time) region-by-region by VerbType in the verb distractor condition (left) and the noun distractor condition (right) in Experiment 2a. Error bars represent standard error of the mean. Gray areas represent the regions of interest for each distractor type.

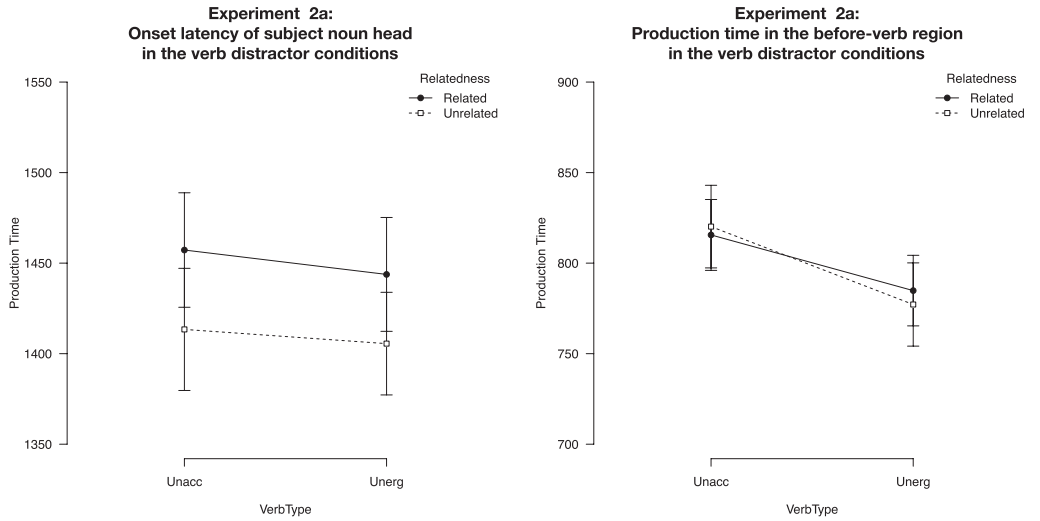

Fig. 8. Production time in two regions of interest in the verb distractor conditions in Experiment 2a. Error bars represent 95% within-subject confidence intervals.

Table 5
Results of mixed effects analyses on the two regions of interest in the verb distractor conditions in Experiment 2a.

	Estimate	SE	z	p
Before-subject noun region				
Intercept	1617	29	54.85	< .001***
Relatedness	−44	23	−1.92	.055
VerbType	−4	27	−0.14	.887
Relatedness × VerbType	5	27	0.19	.846
Before-verb region				
Intercept	883	27	33.20	< .001***
Relatedness	0	9	0.00	.998
VerbType	−28	38	−0.74	.461
Relatedness × VerbType	−19	18	−1.04	.300

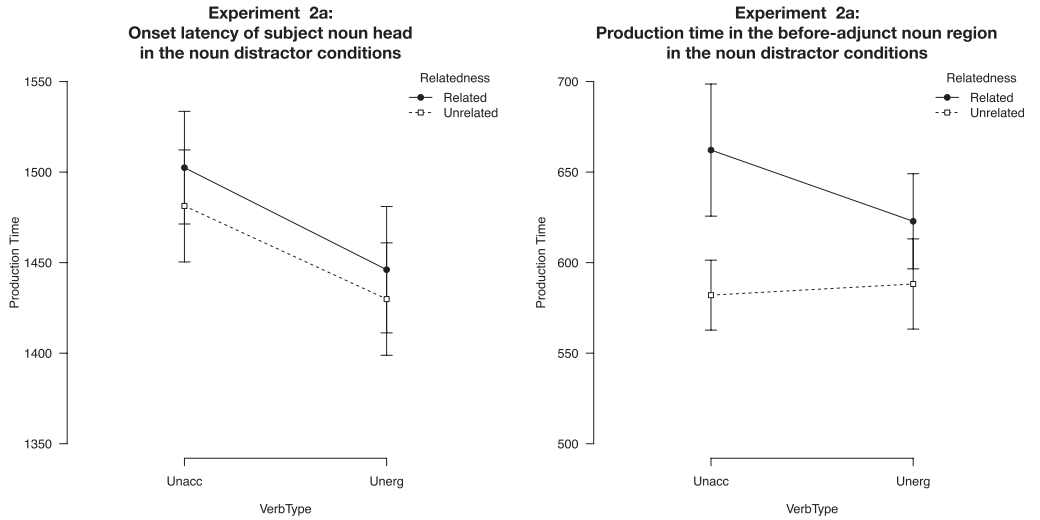

Fig. 9. Production time in two regions of interest in the noun distractor conditions in Experiment 2a. Error bars represent 95% within-subject confidence intervals.

As can be seen in Table 5, the analysis of the before-verb region revealed no significant main effect of Relatedness, VerbType, or the interaction between Relatedness and VerbType (all *ps* > .29). Thus, Experiment 2a provides no evidence regarding when speakers retrieve unergative verbs.

3.2.2. The effect of noun distractors on production time

Next, we report analyses in the noun distractor conditions. Production times in the two regions of interest by each condition are shown in Fig. 9. Analysis on the before-subject noun region of interest yields little evidence that speakers plan the adjunct noun before the articulation onset of the subject noun. As can be seen in Table 6, speakers were not faster or slower across all conditions: the main effect of Relatedness, VerbType and the interaction between Relatedness and VerbType were not significant (all *ps* > .18).

In contrast, the analysis on the before-adjunct noun region revealed that speakers exhibited a semantic interference effect on the adjunct noun as they produced the preposition and the determiner preceding that noun. As can be seen in Table 6 and Fig. 9, the main effect of Relatedness was significant ($\hat{\beta} = -57$, $SE = 15$, $z = -3.81$, $p < .01$). The main effect of VerbType ($p = .19$) or the interaction between Relatedness and VerbType was not significant ($p > .36$), suggesting that speakers spend about the same time between two verb types in this region. There was no indication that the effect of relatedness differed by VerbType.

Given the robust adjunct noun interference effect, we conducted an additional analysis that aimed to characterize the time-course of the interference effect (see Analysis section in Experiment 1 for detail). This analysis revealed, just like in Experiment 1, that the two-way interaction between Relatedness and Region was significant ($\hat{\beta} = 0.14$, $SE = 0.06$, $t = 2.15$, $p = .03$), suggesting that the interference effect was stronger in the before-adjunct noun region than in the before-subject noun region. The main effect of Relatedness was also significant ($\hat{\beta} = -0.13$, $SE = 0.03$, $t = -4.10$, $p < .001$), but we avoid interpreting this main effect due to the presence of the higher-order interaction. No other effects were significant (*ps* > .10).

3.3. Results of Experiment 2b (SOA = 300 ms)

10.89% of all data points (531 out of 4896 trials) were identified as erroneous trials, and so were excluded from subsequent

Table 6
Results of the mixed effects model analyses on the two regions of interest in the noun distractor conditions in Experiment 2a.

	Estimate	SE	z	p
Before-subject noun region				
Intercept	1632	36	45.81	< .001***
Relatedness	−22	24	−0.91	.363
VerbType	−47	36	−1.31	.189
Relatedness × VerbType	−14	35	−0.40	.689
Before-adjunct noun region				
Intercept	717	27	27.84	< .001***
Relatedness	−57	15	−3.81	.0001***
VerbType	−23	32	−0.71	.476
Relatedness × VerbType	15	16	0.92	.357

Table 7
The by-subject mean production time in Experiment 2b, along with the accuracy (in %), by each condition.

DistractorType	VerbType	Relatedness	Onset	The	octopus	below	the	spoon	is	accuracy %
Verb	Unacc	Related	1144	180	511	391	158	543	278	88
Verb	Unacc	Unrelated	1156	195	517	407	166	535	264	91
Verb	Unerg	Related	1129	190	507	387	151	531	243	89
Verb	Unerg	Unrelated	1122	190	499	379	147	524	240	90
Noun	Unacc	Related	1165	206	521	431	194	561	282	87
Noun	Unacc	Unrelated	1167	186	526	389	158	528	261	89
Noun	Unerg	Related	1151	189	504	423	189	571	254	87
Noun	Unerg	Unrelated	1120	188	501	398	154	532	235	92

analyses. The word-by-word mean production time, along with the accuracy rate (in %) by each condition is summarized in Table 7. Fig. 10 is a difference plot showing the semantic interference effect on each word's production time.

3.3.1. The effect of verb distractors on production time

First, we report the analysis of the verb distractor conditions. As can be seen in Table 8 and Fig. 11, this analysis provides no evidence that speakers experienced semantic interference effects before they started to speak the subject noun. The main effect of Relatedness, VerbType, or the interaction between Relatedness and VerbType were not significant (all *ps* > .40).

As can be seen in Table 8, the analysis of the before-verb region revealed no significant main effect of Relatedness, VerbType, or the interaction between Relatedness and VerbType (all *ps* > .16). Thus, these data do not provide any insight as when speakers retrieve unergative verbs.

3.3.2. The effect of noun distractors on production time

Next, we report the analysis on the noun distractor conditions. The production times in the two regions of interest by each condition are shown in Fig. 12. The analysis on the before-subject noun region of interest provides no strong evidence that speakers retrieve the adjunct noun before the articulation onset of the subject noun. The main effect of Relatedness, VerbType, and the interaction between Relatedness and VerbType were not significant (*ps* > .22).

In contrast, the analysis on the before-adjunct noun region revealed that speakers experienced semantic interference effect on the adjunct noun as they produced the preposition and the determiner. As can be seen in Table 9 and Fig. 12, the main effect of Relatedness was significant ($\hat{\beta} = -105$, $SE = 17$, $t = -6.07$, $p < .001$). The main effect of VerbType was not significant (all *ps* > .95), suggesting that speakers spend about the same time between two verb types in this region. There was no indication that the effect of relatedness differed by VerbType: The interaction between Relatedness and VerbType was not significant ($p = .55$).

Given the robust adjunct noun interference effect, we conducted an additional analysis that aimed to characterize the time-course of the interference effect (see Analysis section in Experiment 1 for detail). This analysis revealed, just like in Experiment 1 and in Experiment 2b, that the two-way interaction between Relatedness and Region was significant ($\hat{\beta} = 0.13$, $SE = 0.06$, $t = 2.22$, $p = .03$), suggesting that the interference effect was stronger in the before-adjunct noun region than in the before-subject noun

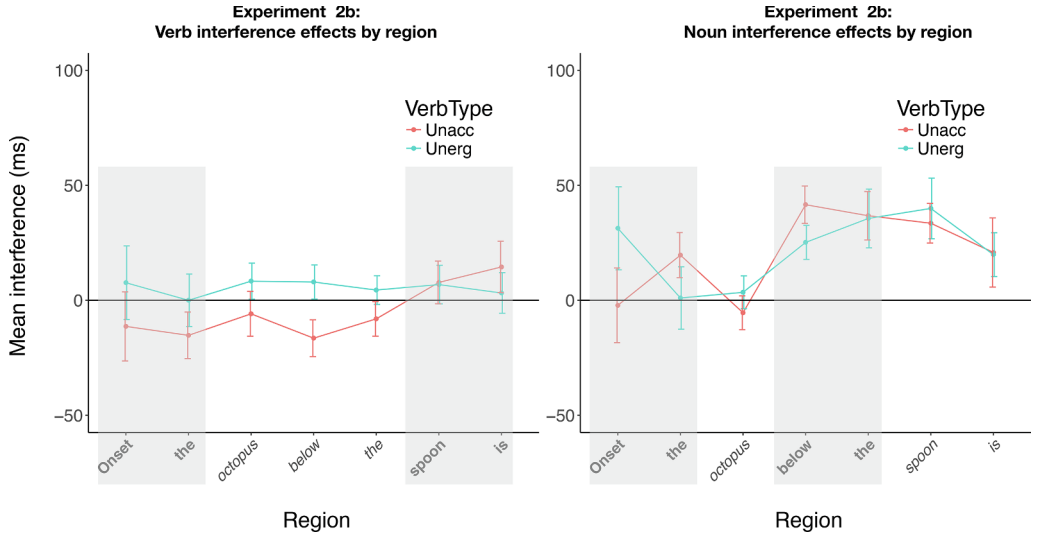

Fig. 10. Interference effect (in raw production time) on region-by-region production time by VerbType in the verb distractor condition (left) and the noun distractor condition (right) in Experiment 2b. Error bars represent standard error the of mean. Gray areas represent the regions of interest for each distractor type.

Table 8
Results of mixed effects model analyses on the two regions of interest in the verb distractor conditions in Experiment 2b.

	Estimate	SE	z	p
Before-subject noun region				
Intercept	1497	32	46.38	< .001***
Relatedness	−4	11	−0.32	.748
VerbType	−27	32	−0.83	.409
Relatedness × VerbType	−15	21	−0.70	.484
Before-verb region				
Intercept	864	26	33.68	< .001***
Relatedness	−16	12	−1.39	.166
VerbType	−41	30	−1.37	.169
Relatedness × VerbType	12	20	0.64	.525

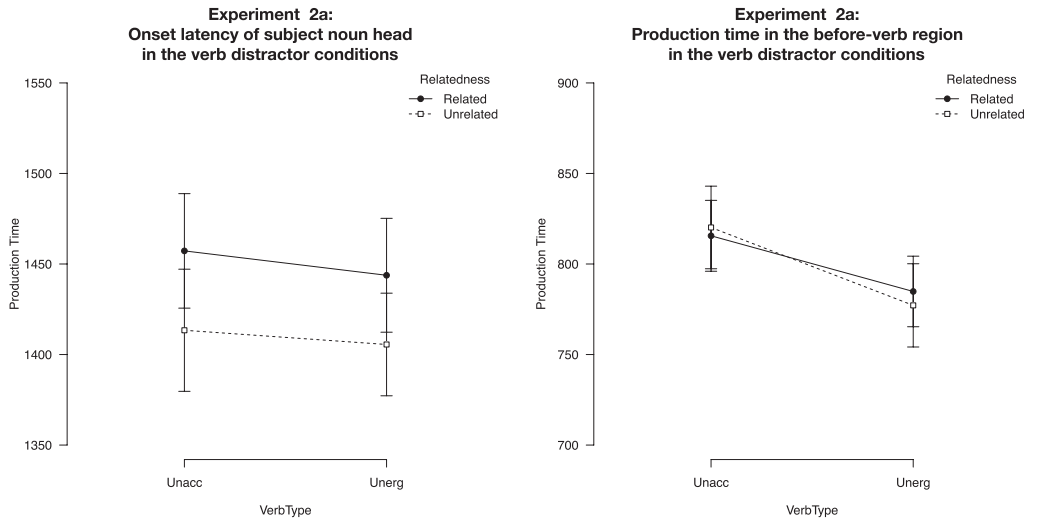

Fig. 11. Production time in two regions of interest in the verb distractor conditions in Experiment 2b. Error bars represent 95% within-subject confidence intervals.

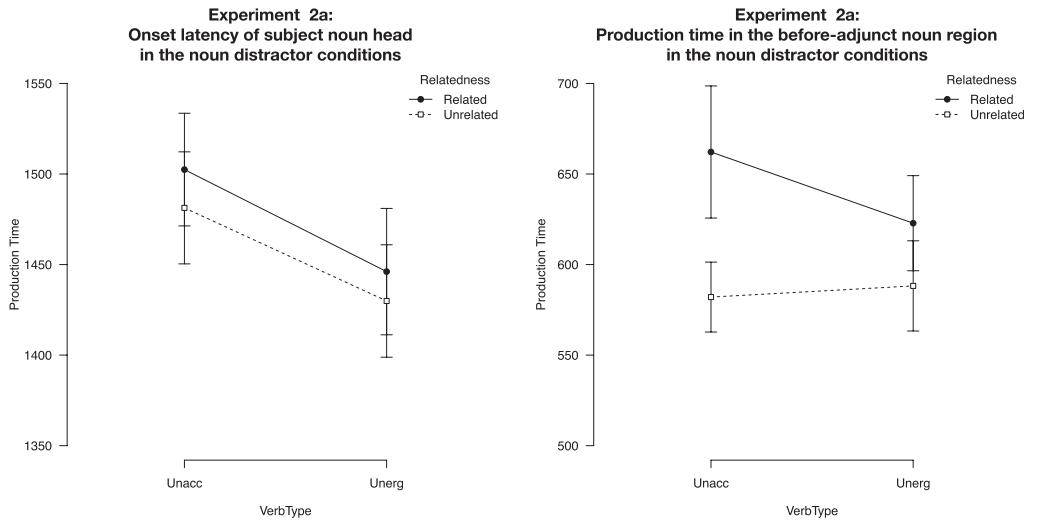

Fig. 12. Production time in two regions of interest in the noun distractor conditions in Experiment 2b. Error bars represent 95% within-subject confidence intervals.

Table 9
Result of mixed effects model analyses on the two regions of interest in the noun distractor conditions in Experiment 2b.

	Estimate	SE	z	p
Before-subject noun region				
Intercept	1489	39	37.81	< .001***
Relatedness	−7	12	−0.60	.55
VerbType	−37	49	−0.76	.447
Relatedness × VerbType	−31	25	−1.21	.226
Before-adjunct noun region				
Intercept	661	24	28.04	< .001***
Relatedness	−105	17	−6.07	< .001***
VerbType	−2	35	−0.07	.946
Relatedness × VerbType	21	35	0.59	.554

region. The main effect of Relatedness was also significant ($\hat{\beta} = -0.13, SE = 0.03, t = -4.43, p < .001$), but we avoid interpreting this main effect due to the presence of the higher-order interaction. No other effects were significant ($ps > .32$).

3.4. Discussion

Experiment 2a and 2b yielded two main results. First, Experiments 2a and 2b both independently replicated the adjunct noun interference effect found in the total production time of the preposition and the second determiner preceding that noun. This result is a replication of Experiment 1 and suggests that under these task conditions, speakers retrieve adjunct nouns on a just-in-time basis. Experiment 2a, like Experiment 1, did not show any adjunct noun interference effect before the subject noun onset.

Second, as in Experiment 1, we again failed to find semantic inference effects on unergative verbs in either region of interest (though there were numerical differences in the predicted direction in several production time measures, as can be seen in Table 4). Thus, there remain two possible reasons for the absence of unergative verb interference effects. It is possible that the lack of unergative verb interference was due to the relatively long time-lag between the distractor presentation and the timing of unergative verb retrieval. Under this explanation, the distractor presented with the SOA of 0 ms or 300 ms was still not late enough to interfere with the unergative verb-retrieval process that was hypothesized to happen late in an utterance. This interpretation is consistent with the hypothesis that speakers plan unergative verbs on a just-in-time basis. It is, however, also possible that the absence of the effect was because the selected distractors were not sufficiently related to the unergative verbs to-be-produced. It is also possible that individual speakers vary in when they retrieve unergative verbs, so no single region showed the consistent unergative verb interference effect. We continue to evaluate these possibilities in the subsequent experiments.

4. Experiment 3

So far, the unaccusative verb interference effect on the subject noun onset latency was only observed when the distractor words were presented 150 ms before the picture, that is, only in Experiment 1. This result is not entirely surprising given previous studies that show that the semantic interference effect is sensitive to SOA manipulations (Schriefers et al., 1990). However, to be more confident that the difference was indeed due to the SOA difference, we tested whether the unaccusative semantic interference effect in Experiment 1 is replicable, but potentially only when the distractor was presented early (i.e., with an SOA of −150 ms). Thus, in Experiment 3, we used the same pictures and the same set of verb distractors (with the same paring between pictures and distractors) as in Experiments 1, 2a and 2b but manipulated SOA as a within-subject factor. The primary goal was to examine whether the critical effect (Relatedness × VerbType interaction on a subject noun onset latency) replicates, potentially only with the SOA of −150 ms.

Experiment 3 focused mainly on verb planning, so we used verb distractors only. This decision was made because the effects associated with noun planning were unambiguously present three times in Experiments 1, 2a, and 2b.

4.1. Methods

4.1.1. Participants

Sixty undergraduate students at the University of California, San Diego participated in the experiment for course credit. Two participants were replaced because of their poor accuracy rate (< 50% in at least one of the conditions). All participants reported that they learned English as their first language. None participated in Experiment 1, 2a or 2b. Informed consent was obtained for each participant before the experimental session.

4.1.2. Materials, procedure and analysis

The same materials, procedures, and analyses were used as in Experiments 1, 2a, and 2b, except that all noun distractors were replaced with verb distractors. This means that there was no DistractorType factor, and all verb distractors were used twice in the related condition (once at each SOA) and twice in the unrelated condition (once at each SOA). Also, Experiment 3 manipulated SOA as a within-subject factor (−150 ms vs. 0 ms).

Table 10
The by-subject mean production time by each condition by each region in Experiment 3.

SOA	VerbType	Relatedness	Onset	The	octopus	below	the	spoon	is	accuracy %
−150 ms	Unacc	Related	1167	204	535	409	164	554	288	87
−150 ms	Unacc	Unrelated	1110	180	530	401	146	549	265	88
−150 ms	Unerg	Related	1101	183	506	400	172	538	275	88
−150 ms	Unerg	Unrelated	1113	185	499	398	146	533	257	90
0 ms	Unacc	Related	1156	207	525	406	156	545	268	86
0 ms	Unacc	Unrelated	1149	194	532	410	161	546	265	89
0 ms	Unerg	Related	1118	188	516	414	159	542	273	87
0 ms	Unerg	Unrelated	1142	197	515	405	174	536	265	89

Also, we analyzed each level of the SOA factor separately, for both theoretical and practical reasons. First, our central claims do not hinge on whether relatedness effects vary reliably across different SOAs. Second, even if we failed to find the three-way interaction between Relatedness, VerbType and SOAs, it could merely reflect lack of statistical power to detect the three-way interaction. Third, fitting Inverse Gaussian distribution on the overall data involves convergence problems that are difficult to resolve.

4.2. Results

12.07% of all data points (695 out of 5760 trials) were excluded from subsequent analyses as erroneous trials. The word-by-word mean production time, along with the accuracy rate (in %) in each condition are summarized in Table 10. Fig. 13 is the difference plot showing the semantic interference effect for each word.

In the −150 SOA conditions, speakers were around 81 ms slower to start speaking the subject noun given related distractors in the unaccusative condition. In comparison, speakers were 14 ms faster to start speaking the subject head noun given related distractors in the unergative condition. This pattern is shown in Fig. 14. As can be seen in Table 11, there was a significant two-way interaction between Relatedness and VerbType in the −150 SOA condition ($\hat{\beta}$ = 77, SE = 22, z = 3.52, p < .001). The subsequent pairwise comparison revealed that the effect of Relatedness was significant in the unaccusative conditions ($\hat{\beta}$ = −62, SE = 16, z = 3.72, p < .001), but not in the unergative conditions ($\hat{\beta}$ = 15, SE = 17, z = 0.92, p = .36). The main effect of Relatedness was marginally significant ($\hat{\beta}$ = −24, SE = 13, z = −1.88, p = .06), but we avoid interpreting this marginally significant effect in the presence of the interaction. The main effect of VerbType was not significant (p = .69).

In addition, in the before-verb region, speakers were 26 ms slower to speak the adjunct noun and the auxiliary verb. We found a significant main effect of Relatedness in this region ($\hat{\beta}$ =-28, SE = 9, z = −3.22, p = .001). The subsequent pairwise comparison revealed that the interference effect was present in both the unaccusative ($\hat{\beta}$ =-28, SE = 12, z = −2.31, p = .02) and the unergative conditions ($\hat{\beta}$ =-28, SE = 10, z = −2.56, p = .01). There was no evidence that this interference effect differed by VerbType; the interaction between Relatedness and VerbType was not significant (p = .99). The main effect of VerbType was not significant (p = .54).

Given the robust unaccusative verb interference effect in the before- subject noun region in the −150 SOA condition, we

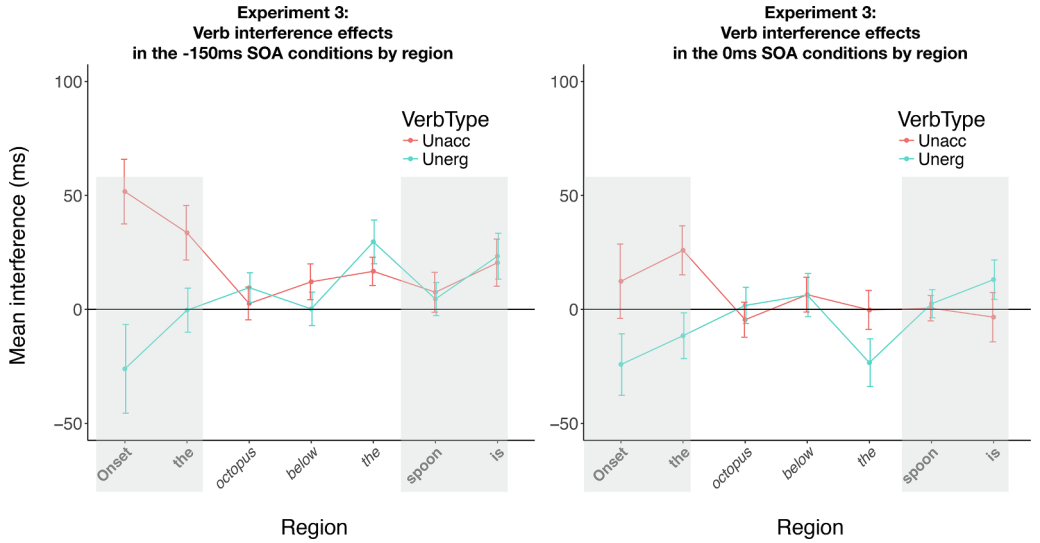

Fig. 13. Interference effect (in raw production time) by each region by VerbType in the −150 ms SOA condition (left) and in the 0 ms SOA condition (right) in Experiment 3. Error bars represent standard errors. The regions of interest are shaded in gray.

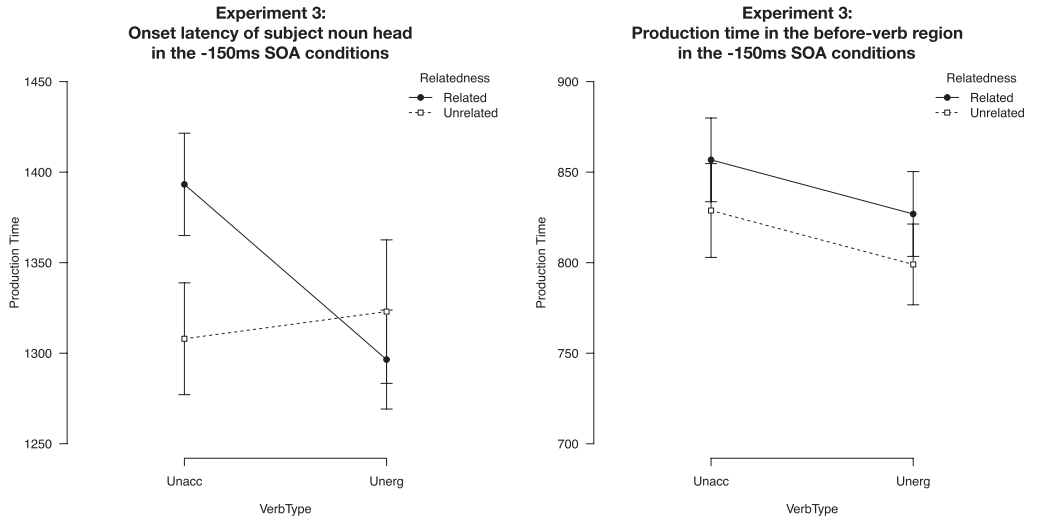

Fig. 14. Production time in two regions of interest in the −150 ms SOA conditions in Experiment 3. Error bars represent 95% within-subject confidence intervals.

Table 11
Results of mixed effects model analyses on the before-subject noun region in the negative SOA conditions in Experiment 3.

	Estimate	SE	z	p
Before-subject noun region				
Intercept	1441	39	36.76	< .001***
Relatedness	−24	13	−1.88	.060
VerbType	−30	77	−0.39	.694
Relatedness × VerbType	77	22	3.52	< .001***
Before-verb region				
Intercept	921	33	28.32	< .001***
Relatedness	−28	9	−3.22	.001**
VerbType	−24	38	−0.60	.548
Relatedness × VerbType	−0	15	−0.02	.988

conducted an additional analysis that aimed to characterize the time-course of the unaccusative interference effect. This analysis revealed that the two-way interaction between Relatedness and Region was marginally significant ($\hat{\beta} = -0.13$, $SE = 0.07$, $t = -1.72$, $p = .08$). Though this effect was marginally significant, the same effect was significant in Experiment 1. Thus, overall, the time-course analysis suggests that the interference effect was stronger in the before-subject noun region than in the before-verb region. The main effect of Relatedness was significant ($\hat{\beta} = -0.14$, $SE = 0.03$, $t = -3.67$, $p < .001$), perhaps because in Experiment 3 the unaccusative verb interference effect was present in both the before-subject and before-verb regions. No other effects were significant ($ps > .15$).

In the 0 ms SOA conditions, speakers were 20 ms slower to start speaking the subject noun given the related verb distractors in the unaccusative conditions. In contrast, they were 33 ms faster in the unergative conditions. This pattern is shown in Fig. 15. As can be seen in Table 12, the two-way interaction between Relatedness and VerbType was not significant ($p = .30$). The main effect of VerbType and the interaction between Relatedness and VerbType were not significant ($ps > .40$).

In the before-verb region, we failed to find a reliable semantic interference effect on unergative verbs. As can be seen in Table 11, the main effect of Relatedness, VerbType, and the interaction between Relatedness and VerbType were not significant ($ps > .21$). Thus, we continue to lack evidence from this paradigm as to when speakers retrieve unergative verbs, and it also remains unclear whether the failure to obtain the semantic interference effect on unergative verbs is due to the long time interval between the distractor processing and unergative verb retrievals, due to insufficiently related distractors, or due to the individual variation in when to retrieve unergative verbs.

4.3. Discussion

Experiment 3 replicated the pattern of verb interference effects observed in Experiment 1 and 2. In the −150 SOA condition, speakers exhibited the verb interference effect in the subject noun onset selectively in the unaccusative condition. This result reinforces the conclusion that speakers retrieve unaccusative verbs selectively in advance of the subject head noun articulation onset. The critical interaction effect (that indicates selective interference effects in the unaccusative condition) was not reliable in the 0 ms

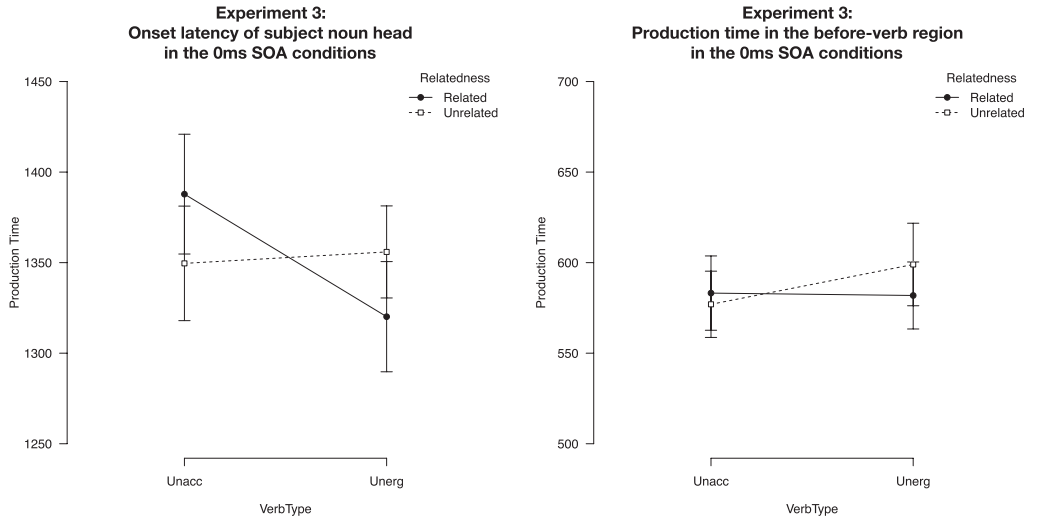

Fig. 15. Production time in two regions of interest in the 0 ms SOA conditions in Experiment 3. Error bars represent 95% within-subject confidence intervals.

Table 12
Results of mixed effects model analyses on the before-subject noun region of interest in the 0 ms SOA conditions in Experiment 3.

	Estimate	SE	z	p
Before-subject noun region				
Intercept	1458	23	62.62	$< .001^{***}$
Relatedness	5	20	0.27	.784
VerbType	−19	23	−0.84	.401
Relatedness × VerbType	27	26	1.03	.304
Before-verb region				
Intercept	924	22	42.28	$< .001^{***}$
Relatedness	−19	16	−1.17	.241
VerbType	−8	24	−0.31	.756
Relatedness × VerbType	−25	20	−1.25	.211

SOA condition. Thus, the absence of the verb interference effect in the unaccusative conditions in Experiment 2a and 2b is likely because distractor words were presented too late to reliably affect the unaccusative verb retrieval process, which by the current hypothesis occurs early in the sentence planning process.

Finally, we found the verb semantic interference effect in the 150 ms SOA condition, but this effect was not selective to the unergative conditions. This effect might suggest that speakers retrieve unergative verbs in the before-verb region, and that speakers retrieve unaccusative verbs *for the second time* in the before-verb region, or that some speakers retrieved unaccusative verb on a just-in-time basis on the subset of the trials. However, note that this effect was not found in the 0 ms SOA condition, or in Experiment 1, 2a, or 2b. Thus, we avoid giving a strong interpretation to this effect. Finally, we found a significant effect of unergative verb interference in the −150 ms SOA condition, but we failed to find the same effect in the 0 ms SOA condition. Thus, the evidence for the unergative interference effect in the before-verb region remains inconclusive. Thus, we continue to investigate the issue of why unergative verb interference effect is absent or weak at the before-verb region in Experiment 4 and 5.

5. Experiment 4

Experiments 1–3 established that speakers exhibit a semantic interference effect before they start speaking the subject noun of unaccusative sentences, but not unergative sentences. However, the experiments fail to offer any information about when unergative verbs are planned. We offered a potential explanation of why the unergative interference effect was not reliably observed: the weak unergative interference may be due to the temporal interval between the timing of distractor presentation and the timing of unergative lemma retrieval processes. But it remains possible that the unergative distractors used in Experiment 1–3 were insufficiently related to the target unergative verbs. The goal of Experiment 4 was to examine this possibility by removing the intervening material between the subject noun phrase and the verbs, as in the following example sentences:

(1) The octopus is boiling [unaccusative]
(2) The octopus is swimming [unergative]

Unlike in Experiment 1–3, the target sentences did not contain adjunct modifiers, but they were otherwise identical to the target sentences in Experiment 1–3. If the absence of the reliable effect of unergative verb interference was due to the relatively long time-lag between the distractor presentation and unergative verb retrieval, a robust semantic interference effect from unergative verb distractors in Experiment 4 is predicted. On the other hand, if the inconsistency of the unergative verb interference effects was merely due to the unergative verb distractors being insufficiently related to those unergative verbs, the unergative verb interference effect should not be observed in Experiment 4.

5.1. Methods

5.1.1. Participants

Fifty undergraduate students at the University of California, San Diego, participated in the experiment for course credit. Three participants were replaced because they did not follow instructions. One more participant was replaced because of their high error rate ($>$ 50% in one of the conditions). All participants reported that they learned English as their first language. None participated in Experiment 1–3. Informed consent was obtained for each participant before the experimental session.

5.1.2. Materials

The same set of event pictures used in Experiment 1–3 were used in Experiment 4, except that the pictures for the adjunct nouns (e.g., *lemon* and *spoon*) were removed (to simplify the task). The same set of verb distractors were used. The pairing between verb distractors and event pictures were also the same as in Experiment 1–3.

5.1.3. Procedure

Participants underwent a similar familiarization as in Experiment 1–3, where they studied the entire set of event pictures with target sentences and then practiced describing each picture once. Following this familiarization phase, they performed the similar sentence description task with distractor words superimposed on each picture (i.e., the ePWI task). Stimulus presentation parameters (e.g., when the fixation cross or the stimulus picture appear and disappear, etc.) were identical to Experiment 1–3. The distractors appeared 150 ms before the presentation of the pictures. (i.e., the SOA was -150 ms).

Because the primary purpose of this experiment was to test whether the related distractors were equally effective between the unaccusative and unergative conditions, we focused on the onset latency of the verb, rather than on the production time of pre-defined regions of interest. This decision was made because previous studies showed that verb interference effects in unergative sentences could appear later than sentence onset (Momma et al., 2018).

5.1.4. Analysis

The same analysis procedures were used as in Experiment 1–3.

5.2. Results

7.8% of all trials (188 out of 2400 trials) were identified as erroneous and thus were excluded from subsequent analyses. Table 13 summarizes the region-by-region production time, along with the accuracy rate in each condition. Fig. 16 is the word-by-word difference plot showing the interference effect. Speakers were on average 98 ms slower to start speaking subject nouns given the related verb distractors. In the before-subject noun region, the main effect of Relatedness was significant ($\hat{\beta} = -76$, $SE = 12$, $z = -6.23$, $p < .001$). There was no evidence that this semantic interference effect was different between the unaccusative and the unergative conditions. There was no interaction between Relatedness and VerbType ($p > .23$), and pairwise comparisons revealed that the interference effect was significant both in the unaccusative ($\hat{\beta} = -62$, $SE = 16$, $z = -3.72$, $p < .001$) and unergative conditions ($\hat{\beta} = -89$, $SE = 17$, $z = -5.35$, $p < .001$). The main effect of VerbType was not significant ($p > .30$).

In the before-verb region (e.g., *octopus is*), speakers were marginally slower to speak: the main effect was Relatedness was marginally significant ($\hat{\beta} = -12$, $SE = 6$, $z = -1.94$, $p = .05$). There was no main effect of VerbType ($p = .46$) or interaction between Relatedness and VerbType ($p = .64$). This marginal effect may suggest that speakers might occasionally retrieve verbs later than the subject noun onset. Indeed, consistent with the hypothesis that speakers retrieve unergative verbs on a just-in-time basis, this marginally significant effect was numerically stronger in the unergative conditions (18 ms) than in the unaccusative conditions (11 ms).

Table 13
The by-subject mean production time by each condition by each region in Experiment 4.

VerbType	Condition	Onset	The	octopus	is	accuracy
Unaccusative	Related	1171	167	488	250	91
Unaccusative	Unrelated	1100	169	486	241	92
Unergative	Related	1133	183	477	238	93
Unergative	Unrelated	1056	161	466	231	93

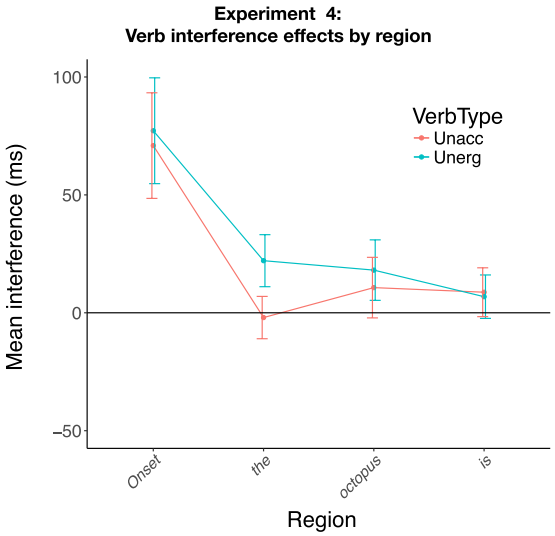

Fig. 16. Interference effect (in raw production time) on region-by-region production time by VerbType in Experiment 4. Error bars represent standard error of the mean.

5.3. Discussion

Experiment 4 showed that distractor verbs in the unergative conditions (as well as in the unaccusative conditions) reliably elicited the verb interference effect (in the subject onset latency measure). If anything, the magnitude of semantic interference effect in the unergative conditions was numerically larger than that in the unaccusative conditions. Thus, the results of Experiment 4 suggests that the lack of unergative interference in Experiments 1–3 was unlikely to be due to the distractors being insufficiently related to the unergative verbs.

However, the result suggests that the timing of unergative verb interference in Experiment 4 is inconsistent with the previous study by Momma et al. (2018). Namely, Momma et al. (2018) observed that the unergative verb interference effect was found in the total production time of the subject noun and the auxiliary verb, not in the onset latency measure. In contrast, the current study showed that speakers experienced a similar amount of semantic interference effect on unergative verbs and on unaccusative verbs in the onset latency measure.

This discrepancy may be explained if speakers can be flexible in terms of how many words they retrieve before starting to speak. It is likely that many factors modulate whether speakers retrieve unergative verbs in advance of speaking their subject head nouns. Among them are working memory load (Wagner, Jescheniak, & Schriefers, 2010), time pressure, task difficulty (Ferreira & Swets, 2002), recent experience (Konopka, 2012), the phonological length of the first word of an utterance (Griffin, 2001) and potentially numerous other factors (e.g., how careful speakers want to be in what they say). The difference between Experiment 4 and Momma et al. (2018) may be due to many differences in how the tasks were set up, the properties of picture stimuli or target sentences, and so forth. One clear difference between Momma et al. (2018) and Experiment 4 is that only Momma et al. (2018) involved a type of stop-signal task (Lappin & Eriksen, 1966; Logan & Cowan, 1984; Vince, 1948; Verbruggen & Logan, 2008; Matzke, Verbruggen, & Logan, 2018), in the sense that participants in Momma et al.'s study were asked to suppress their speech conditionally when they saw a series of x's (xxxx) instead of distractor words. This additional stop-signal task might have increased the task difficulty in Momma et al. (2018), potentially reducing how many words speakers retrieve in advance. To test this hypothesis, we conducted an additional experiment, in which the stop-signal component was added to the main extended picture-word interference task.

6. Experiment 5

Experiment 5 tested whether the introduction of a stop-signal task changes the timing of interference specifically in the unergative conditions. To do so, we adopted the method employed in Momma et al. (2018), in which distractor words are sometimes replaced with series of x's that served as a stop-signal for speech production. In their experiments, as well as in Experiment 5, participants were asked to suppress their speech and remain silent when they see x's instead of real word distractors. This introduced the stop-signal component to the task. In Momma et al. (2018), the reason for doing this was to make sure that speakers could not visually ignore the distractors, but the main question in Experiment 5 was to test whether the introduction of the stop-signal component to the task delays the interference effect observed in Experiment 4 but only in the unergative condition.

6.1. Method

6.1.1. Participant

Fifty-seven undergraduate students at the University of California, San Diego, participated in the experiment for course credit. Five participants were replaced from analyses because they did not follow the instruction. Two more participants were replaced due to high error rate (> 50% in one of the conditions). All participants reported that they learned English as their first language. None participated in Experiment 1–4. Informed consent was obtained for each participant before the experimental session.

6.1.2. Materials

All the pictures, distractors as well as the pairing between pictures and distractors were identical to Experiment 4.

6.1.3. Procedure

The procedure was mostly identical to Experiment 4, except that participants were instructed to suppress their production when the series of x's (xxxx) were presented on the screen instead of distractors words. As a consequence of having trials in which the series of x's were presented, the number of times participants saw each picture was doubled relative to Experiment 4 (from 2 times to 4 times). However, the number of times participants produced the target sentence for each picture was constant.

6.1.4. Analysis

The same analysis procedures as in Experiment 4 were used. However, we additionally analyzed the combined dataset of Experiment 4 and 5, in order to assess the difference between Experiment 4 and 5 due to the introduction of the stop-signal component to the task. In this analysis, we constructed mixed effects models on the production time in the first and second regions of interest (separately), with Relatedness, VerbType, and Experiment (Experiment 4 vs. Experiment 5) as fixed effects, with maximal random effects structures (followed by the simplification of the random effects structure in case of convergence failure according to the same rules as in Experiment 1–4).

6.2. Results

8.1% of all trials (220 out of 2734 trials) were identified as erroneous and thus were excluded from subsequent analyses. Table 14 summarizes the region-by-region production time, along with the accuracy rate in each condition. Fig. 17 is the word-by-word difference plot showing the interference effect. Speakers were overall 82 ms slower to start speaking the subject nouns given the related verb distractors. There was a significant main effect of Relatedness ($\hat{\beta} = -108$, $SE = 26$, $z = -4.11$, $p < .001$). However, the effect of Relatedness was stronger in the unaccusative conditions than in the unergative conditions, as indicated by a significant interaction between Relatedness and VerbType ($\hat{\beta} = 85$, $SE = 38$, $z = 2.27$, $p = .02$). Subsequent pairwise comparisons revealed that the interference effect was significant both in the unaccusative ($\hat{\beta}$ =-151, $SE = 34$, $z = -4.43$, $p < .001$) and in the unergative conditions ($\hat{\beta}$ =-66, $SE = 31$, $z = -2.16$, $p = .03$). The main effect of VerbType was marginally significant ($\hat{\beta}$ =-37, $SE = 22$, $z = -1.72$, $p = .09$), suggesting that speakers were slightly slower to speak the before-verb region in the unaccusative conditions.

In addition, in the before-subject noun region, the cross-experiments analysis (that included the data from both Experiment 4 and 5, n = 107) revealed a three-way interaction between Relatedness, VerbType, and Experiment ($\hat{\beta} = 69$, $SE = 23$, $z = 2.95$, $p = .003$), suggesting that the verb interference effect was stronger in the unaccusative than in the unergative conditions, but only in Experiment 5. Speakers were also overall slower in Experiment 5, perhaps due to the fact that Experiment 5 involved the stop-signal task; the main effect Experiment was significant ($\hat{\beta} = 132$, $SE = 23$, $z = 5.61$, $p < .001$). There was also a significant main effect of Relatedness ($\hat{\beta} = -81$, $SE = 11$, $z = -7.67$, $p < .001$), VerbType ($\hat{\beta} = -30$ $SE = 15$, $z = -1.96$, $p = .048$), and two-way interactions between Relatedness and VerbType ($\hat{\beta} = 37$, $SE = 17$, $z = 2.11$, $p = .03$) and VerbType and Experiment ($\hat{\beta} = -28$, $SE = 14$, $z = -2.08$, $p = .04$). However, we avoid interpreting these effects because of the existence of the higher order three way interaction effect. The effect of the interaction between Relatedness and Experiment was not significant ($p = .81$).

In the before-verb region (e.g., *octopus is*), speakers were not overall slower to utter the subject noun and auxiliary verb. The main effect of Relatedness was not significant ($\hat{\beta}$ =-8, $SE = 6$, $z = -1.29$, $p > .19$). However, there was an interaction between Relatedness and VerbType ($\hat{\beta} = -38$, $SE = 12$, $z = -3.10$, $p = .002$). Subsequent pairwise comparisons revealed that speakers were slower to speak the before-verb region in the unergative conditions ($\hat{\beta} = -27$, $SE = 8$, $z = -3.24$, $p = .001$) but not in the unaccusative conditions ($\hat{\beta} = 10$, $SE = 9$, $z = 1.20$, $p = .23$). There was no main effect of VerbType ($p > .35$).

In addition, in the before-verb region, the cross-experiment analysis revealed a three-way interaction between Relatedness,

Table 14
The by-subject mean production time by each condition by each region in Experiment 5.

VerbType	Condition	Onset	The	octopus	is	accuracy
Unaccusative	Related	1348	163	484	255	91
Unaccusative	Unrelated	1247	152	497	260	93
Unergative	Related	1258	162	483	234	91
Unergative	Unrelated	1221	147	466	225	92

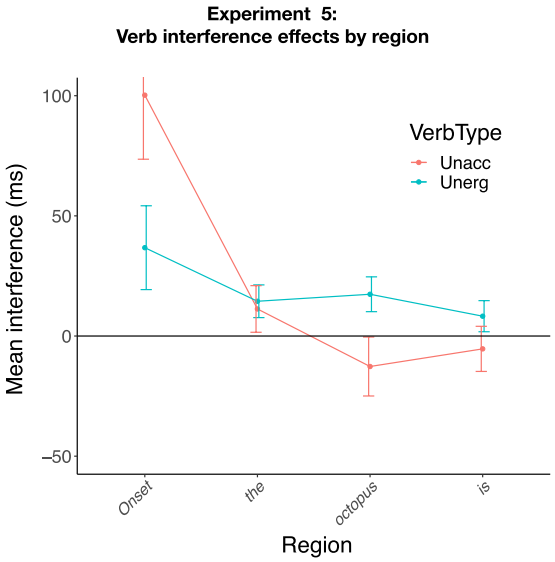

Fig. 17. Interference effect (in raw production time) on region-by-region production time by VerbType in Experiment 5. Error bars represent standard error of the mean.

VerbType, and Experiment ($\hat{\beta} = -30$, $SE = 13$, $z = -2.28$, $p = .02$). This interaction effect suggests that speakers were slower to utter the before-verb region, but only in unergative conditions and only in Experiment 5. The two-way interaction between Relatedness and VerbType was also significant ($\hat{\beta} = -36$, $SE = 14$, $z = -2.61$, $p = .009$), but we avoid interpreting this effect due to the presence of higher-order interaction. Other effects are all non-significant (all $ps > .30$), except that the main effect of VerbType was marginally significant ($\hat{\beta} = -39$, $SE = 20$, $z = -1.93$, $p = .053$).

6.3. Discussion

In Experiment 5, the verb interference effect in the before-subject noun region - the sign of advance verb planning - was stronger in the unaccusative condition than in the unergative condition, as evidenced by the significant two-way interaction between Relatedness and VerbType. Furthermore, this pattern was selective to Experiment 5; the cross-experiment analysis revealed that there was a three-way interaction between Relatedness, VerbType and Experiment. This pattern of results suggests that speakers retrieve unaccusative verbs more consistently before the subject noun onset than unergative verbs, when they engage in the secondary stop-signal task (as in Experiment 5). Although the verb interference effect was significant also within the unergative condition both in Experiment 4 and 5 (as revealed by the pairwise comparisons) unlike in Momma et al. (2018), the results of Experiment 5 are more comparable to Momma et al. (2018). Both Experiment 5 and Momma et al. (2018) showed the stronger interference effect in production of sentences headed by the unaccusative verbs than those headed by the unergative verbs. Moreover, just like in Momma et al. (2018), speakers in Experiment 5 were slower to utter the before-verb region given a related verb distractor, but selectively in the unergative conditions. This timing of verb interference in unergative condition suggests that speakers retrieve unergative verbs on a just-in-time basis, at least in some trials. Taken together, these results show that speakers retrieve unaccusative verbs more consistently in advance than unergative verbs, which can be retrieved on a just-in-time basis.

Because the primary difference between Experiment 4 and 5 was that only Experiment 5 had the additional stop-signal task, it can be argued that when speakers retrieve an unergative verb partially depends on the task structure. This is consistent with the idea that how many words speakers retrieve in advance to is to some extent flexible, in accordance previous claims in the literature (Wagner et al., 2010; Ferreira & Swets, 2002; Konopka, 2012; Levelt, 1989). Importantly, however, the flexibility in whether to retrieve verbs in advance is selective to when speakers produce sentences headed by unergative verbs. When unergative verbs appear relatively far downstream, as in Experiment 1–3, or when the experimental task is relatively complex, as in Experiment 5, speakers tend to retrieve unergative verbs relatively late. In comparison, when unergative verbs appear closer to the sentence onset and the task is relatively simple (as in Experiment 4), speakers can retrieve unergative verbs before sentence onset. This difference follows naturally from the claim that speakers retrieve unaccusative verbs before their subject noun to establish the argument-verb dependency, which exists regardless of how complex the task is and how far downstream the verb appears in linear distance.

The difference between Experiment 4 and Experiment 5 in terms of the timing of verb interference effects is consistent with the idea that how many words speakers retrieve in advance depends on availability of working memory resources (Wagner et al., 2010). Executive control ability, specifically the inhibition component of it (Miyake et al., 2000), is likely to be involved in performing a stop-signal task (Verbruggen & Logan, 2008). And executive control ability is critically involved in working memory (e.g., Baddeley, 2010; Miyake et al., 2000). Thus, when speakers have to perform stop-signal task on top of the picture-word interference task, they might have less working memory resources available for sentence planning, and hence may be less able to plan ahead words that

appear later in an utterance.

Despite the difference in the timing of interference effects between Experiment 4 and 5, these experiments both show that the verb distractors we used for unaccusative and unergative conditions are highly effective in eliciting semantic interference effect. Thus, it is likely that the absence of unergative interference effect in Experiment 1–3 is due to the long time-lag between the distractor presentation and the timing of unergative verb retrieval, due to the individual variation in when speakers retrieve unergative verbs, or both.

7. Individual differences in the timing of unergative verb retrievals

Experiment 4 and 5 both suggest that the distractor verbs we used are effective in eliciting the verb interference effect, so it is unlikely that the absence of unergative verb interference effect in Experiment 1–3 was due to the ineffectiveness of unergative verb distractors. A remaining possibility is the long time-lag from the distractor presentation to the unergative verb retrieval process (the time-lag hypothesis), and as noted above, that the timing of unergative verb retrieval is relatively variable across individuals. By hypothesis, unergative verbs, unlike unaccusative verbs, do not need to be retrieved before their subject arguments. If so, individual speakers can vary in terms of when they retrieve unergative verbs, and because unergative verbs appear at the very end of a sentence this variability is greater than, for example, retrieval of words that appear in the middle of a sentence (e.g., adjunct nouns in the current experiments). If such individual differences exist, it is not surprising that we failed to find anreliable unergative verb interference effect in any single region in an utterance.

To address this issue, we conducted an exploratory analysis that aimed to account for the potential individual differences regarding when speakers retrieve unergative verbs. In this analysis, to estimate when each speaker retrieve unergative verbs, the verb frequency effect in the unergative condition was used as a proxy for estimating when speakers retrieve unergative verbs. It is well known that more frequent words are named faster in picture naming tasks (Oldfield & Wingfield, 1965). The frequency effect is argued to primarily arise at the phonological encoding level (Jescheniak & Levelt, 1994, but see, e.g., Kittredge, Dell, Verkuilen, & Schwartz, 2008). However, we contend that the lemma retrieval process and word form encoding process for unergative verbs (but not necessarily for unaccusative verbs) should occur temporally closely, because there is no theoretical reason to believe that speakers retrieve unergative verb lemmas far in advance of retrieving their word form. If this is the case, the timing of unergative verb frequency effects should correspond roughly to when speakers retrieve unergative verb lemmas even during sentence production. Based on this reasoning, to estimate when each individual speaker tends to retrieve unergative verbs separately for each speakers, we computed the correlation coefficient (Pearson's r) between the production time of each region and the frequency of the (upcoming) verb (obtained from SUBTLEX US corpus, Brysbaert & New, 2009) in the unergative conditions. We then identified the region that exhibited the strongest negative correlation within each individual speaker. The average Pearson's r across speakers in this individually-defined region was −.32. The histogram of the position of the identified region across speakers is plotted in Fig. 18. As can be seen in the histogram, the region that showed the strongest negative correlation most frequently across speakers was the adjunct noun region (e.g., *spoon*), but there are speakers who showed the strongest negative correlation at different regions. Around 80% of speakers (175 out of 219 speakers) showed the strongest negative correlation between verb frequency and production time in the preposition region or later regions, consistent with the claim that speakers retrieve unergative verbs relatively late in an utterance.

The production time of the identified region for each participant was z-score transformed within that participant within that region, to account for scale differences (e.g., the production time of the adjunct noun region is much larger than the production time of the auxiliary verb region). This z-score transformed production time was then analyzed using Gaussian mixed effects modeling. Here, unlike in the other analyses, we could not use the Inverse Gaussian analysis due to the necessity of the z-score transformation. Because z-scores can be negative but the Inverse Gaussian distribution can only accommodate positive values, the Inverse-Gaussian models cannot be applied to this data. This mixed effects model was constructed using the data from unergative conditions only, and included Relatedness, SOA, Experiment, and the interaction between Relatedness and SOA as fixed effects (although the factor Experiment was somewhat redundant with SOA so some contrasts were automatically dropped). The factor Experiment was included as a control predictor, to account for any potential difference across experiments. The SOA was difference-coded, so that the first contrast compared the −150 ms and 0 ms SOA levels, and the second contrast compared the average of −150 ms and 0 ms and 300 ms SOA levels. Experiment was dummy coded, so that Experiment 2a, 2b, and 3 were individually compared against Experiment 1.

Fig. 19 shows the unergative interference effect across three different SOAs, and Table 15 shows the results of the mixed effects analysis. The analysis revealed a highly significant main effect of Relatedness ($\hat{\beta} = -0.09$, $SE = 0.02$, $z = -3.63$, $p = .003$), suggesting that speakers did experience the unergative verb interference effect. There was also a marginally significant interaction between Relatedness and the second contrast of SOA (the average of −150 ms and 0 ms vs. 300 ms, $\hat{\beta} = -0.15$, $SE = 0.08$, $z = -1.90$, $p = .06$), suggesting that the unergative verb interference effect was stronger when the distractor was presented with 300 ms SOA. There was also a marginally significant effect of Experiment that contrasted Experiment 1 and 3, but it was not of our theoretical interest.

Thus, cross-experiments analysis suggests that speakers indeed showed an unergative interference effect but with different timing across different individuals, and that the effect was likely stronger with larger SOAs (as suggested by the marginally significant interaction between Relatedness and SOA). From these results, we argue that the apparent absence of the unergative verb interference in the standard analysis was due to the combination of two factors: (a) individual differences in when speakers retrieve unergative verbs and (b) the relatively long time-lag between the distractor presentation and the unergative verb retrieval process

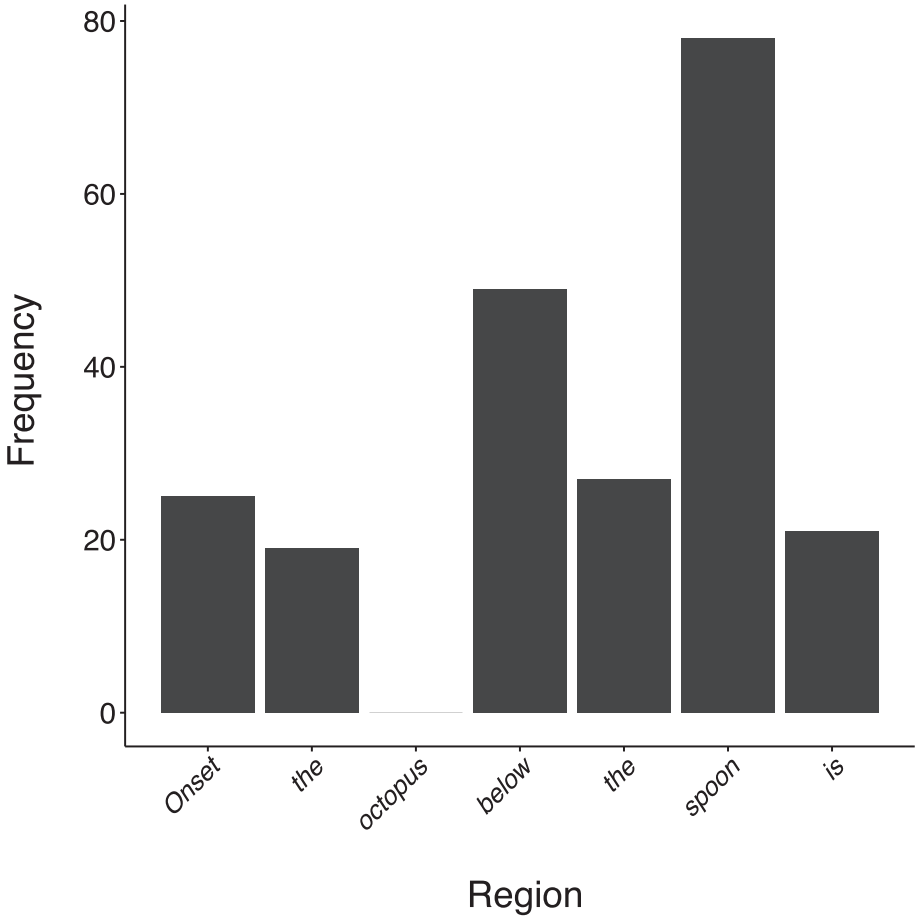

Fig. 18. The histogram of speakers that showed the strongest negative correlation between verb frequency and production time at different regions (total n = 219).

which by our hypothesis occurs relatively late in an utterance. We should note, however, that the analysis reported here is exploratory, so future confirmatory studies are valuable in increasing the confidence in these claims and the analysis methods we reported here.

8. General discussion

We reported six extended picture-word interference studies that examined the time-course of lexical retrieval in sentences headed by two different types of intransitive verbs: unaccusative verbs and unergative verbs. Experiment 1 suggested that speakers reliably retrieved sentence-final unaccusative verbs before the onset of sentence-initial subject nouns, but retrieved sentence-medial adjunct nouns on a just-in-time basis, at least in the majority of trials. Experiments 2a and 2b both supported the conclusion of Experiment 1 that speakers retrieved sentence-medial adjunct nouns on a just-in-time basis. Experiment 3 replicated Experiment 1, again showing the unaccusative verb interference effects in the subject noun onset. Experiment 3 also showed that this interference effect could be more reliably obtained when the distractor is presented sufficiently early. These results are summarized in Figs. 20 and 21. Fig. 20 shows the verb interference effect across three different SOA conditions, and Fig. 21 shows the noun interference effect across three different SOA conditions. Experiment 4 showed that the absence of the reliable semantic interference effect on unergative verbs in Experiment 1–3 was unlikely to be due the distractor verbs being insufficiently related to the unergative verbs. Experiment 4 also showed that unergative verbs *can* be retrieved in advance of the sentence onset when they appear closer to the sentence-initial position, but Experiment 5 showed that the timing of unergative verb retrieval, but not unaccusative verb retrieval, was affected by the secondary stop-signal task; speakers retrieved unergative verbs less consistently before the subject noun onset with the stop-signal task. These results from Experiment 4 and 5 are summarized in Fig. 22. Finally, the cross-experiments exploratory analysis suggests that the apparent lack of unergative interference effects in Experiment 1–3 was likely because of variable the timing of unergative verb retrieval across individuals and because of the relatively long time-lag between the distractor presentation and unergative verb retrieval.

Overall, these results suggest that speakers retrieve sentence-final unaccusative verbs before starting to produce the subject noun,

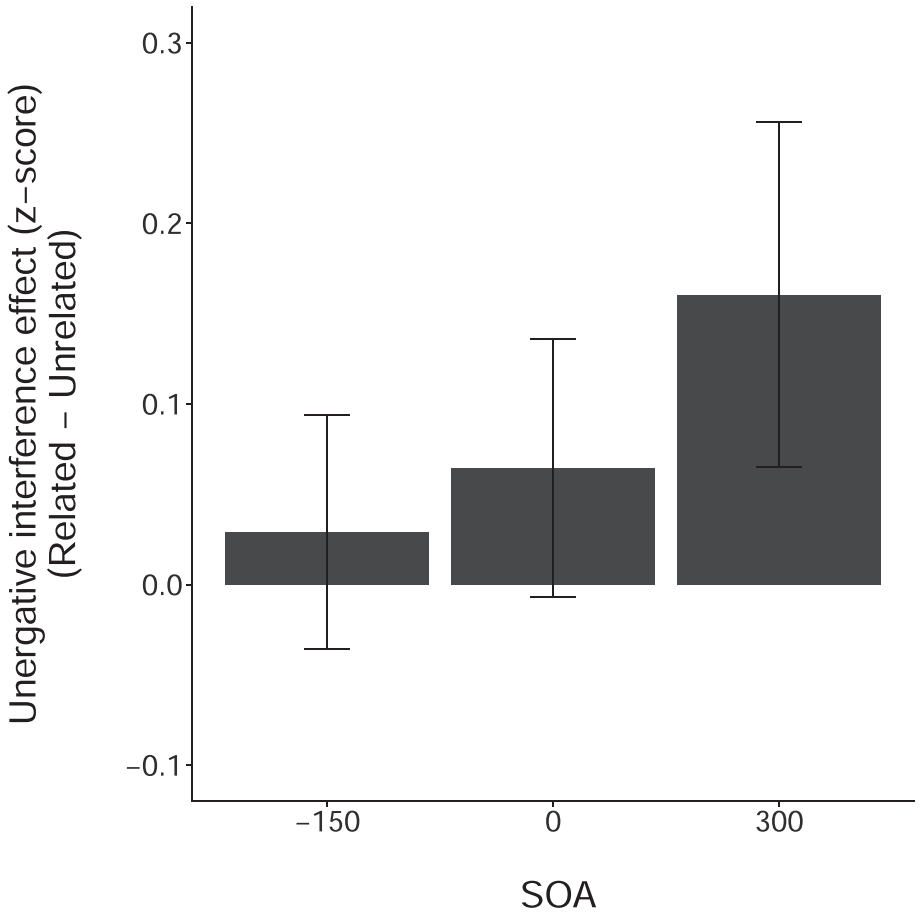

Fig. 19. The mean unergative verb interference effect at the individually defined region (using the unergative verb frequency effect as proxy) by three SOAs. The error bars represent 95% confidence interval.

Table 15
Results of mixed effects model analyses in z-score transformed production time in the region that showed the strongest verb frequency effect for each individual.

	Estimate	SE	z	p
Intercept	−0.07	0.10	0.72	.484
Relatedness	−0.09	0.02	−3.63	.0003***
SOA1 (−150 ms vs. 0 ms)	0.02	0.04	0.67	.505
SOA2 (−150 & 0 ms vs. 300 ms)	−0.02	0.05	−0.42	.672
Experiment1 (Experiment 1 vs. 2a)	−0.04	0.05	−0.08	.444
Experiment2 (Experiment 1 vs. 3)	0.06	0.04	1.76	.08
Relatedness × SOA1	−0.04	0.05	−0.87	.386
Relatedness × SOA2	−0.15	0.08	−1.90	.058

but retrieve sentence-medial adjunct noun and sentence-final unergative verbs more variably, and often later in an utterance. This time-course of lexical retrieval supports the hypothesis that the time-course of sentence planning reflects argument structure above and beyond surface word-order. Thus, we argue that sentence planning is guided by hierarchically defined argument-verb relations, and that the assumption that the order of lexical retrieval transparently reflects surface word order should be dropped.

8.1. The role of argument structure in speaking

To speak grammatically, speakers need to be able to map conceptual representations onto structural representations, according to the grammar of their language. To achieve this mapping, speakers need to access information stored with verbs, because verbs' argument structures determine which event participants map onto which grammatical roles in a manner that is not derivable from conceptual structure alone (Grimshaw, 1990). For this reason, some previous models of grammatical encoding (e.g., De Smedt, 1996;

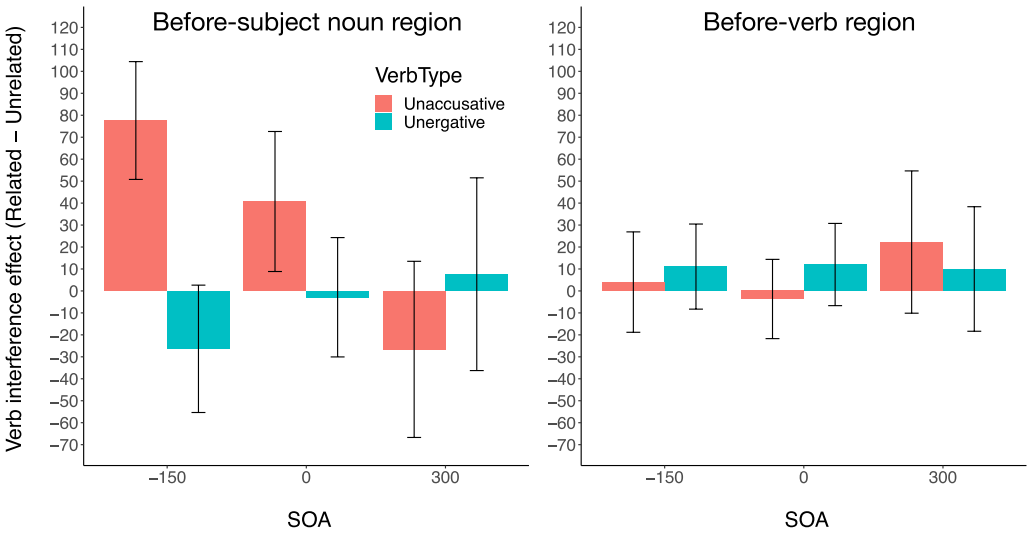

Fig. 20. The semantic interference effects in the verb distractor conditions across three different SOAs in the before-subject noun region (left) and before-verb region (right) in Experiment 1–3. Error bars represent 95% confidence intervals.

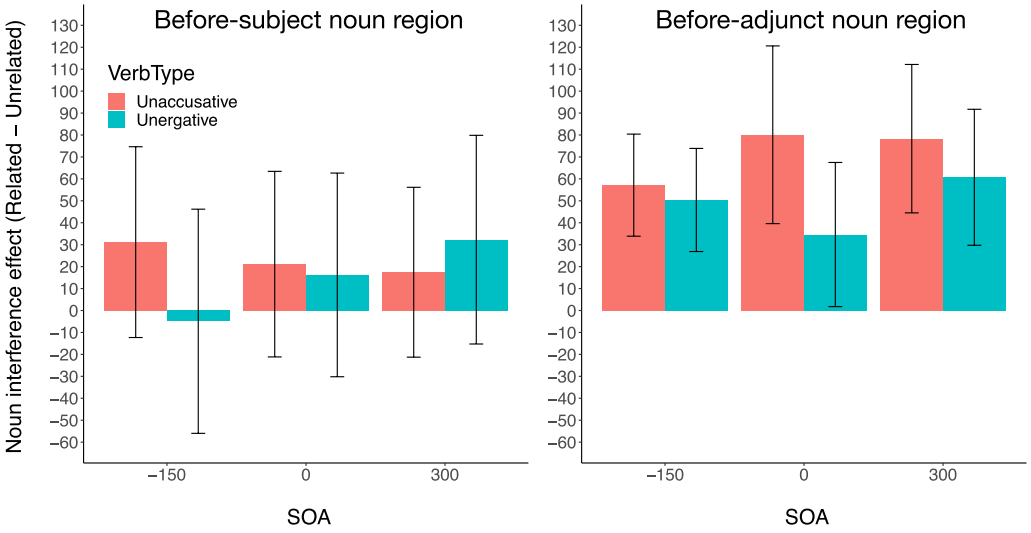

Fig. 21. The semantic interference effects in the noun distractor conditions across three different SOAs in the before-subject noun region (left) and before-adjunct noun region (right) in Experiment 1, 2a, and 2b. Error bars represent 95% confidence intervals.

Ferreira, 2000; Bock & Levelt, 1994) assume verb guidance, the hypothesis that speakers retrieve verbs early in sentence planning and use them to encode the rest of the sentence grammatically. Models that assume verb-guidance normally assume that both subject and object arguments need verb guidance for their grammatical encoding (Bock & Levelt, 2002; Ferreira, 2000). However, the current results, along with the results from previous studies (Momma et al., 2016; Momma et al., 2018), suggest that verb guidance is plausible but selective to internal arguments. Speakers selectively retrieve verbs before the articulation of theme or patient arguments, but not agent arguments. This pattern is naturally congruent with the linguistic hypothesis that internal arguments are the only true arguments of verb roots (Kratzer, 1996; Kratzer, 2003). This view is consistent with the results of previous studies in German (Schriefers et al., 1998) and Japanese (Momma et al., 2016), which both showed that the verb interference effect was absent in the onset latency measure when the first constituent is an agent noun. Thus, a generalization about when speakers retrieve verbs can be stated as: Speakers retrieve verbs before their internal (i.e., non-agentive) arguments.

But the problem with any verb guidance hypothesis is that verbs can appear arbitrarily far from their internal arguments in linear distance. If speakers obligatorily retrieve words sequentially, they must either (a) give up retrieving verbs in advance when they appear further away in a sentence, and instead speak with uncertainty about clausal structure, or (b) plan the entire portion of sentences up to the verbs (see *the head principle* by Martin & Freedman, 2001 for a related view). However, once the assumption that lexical retrieval processes obligatorily occur sequentially is abandoned, speakers can retrieve only the words that are necessary for

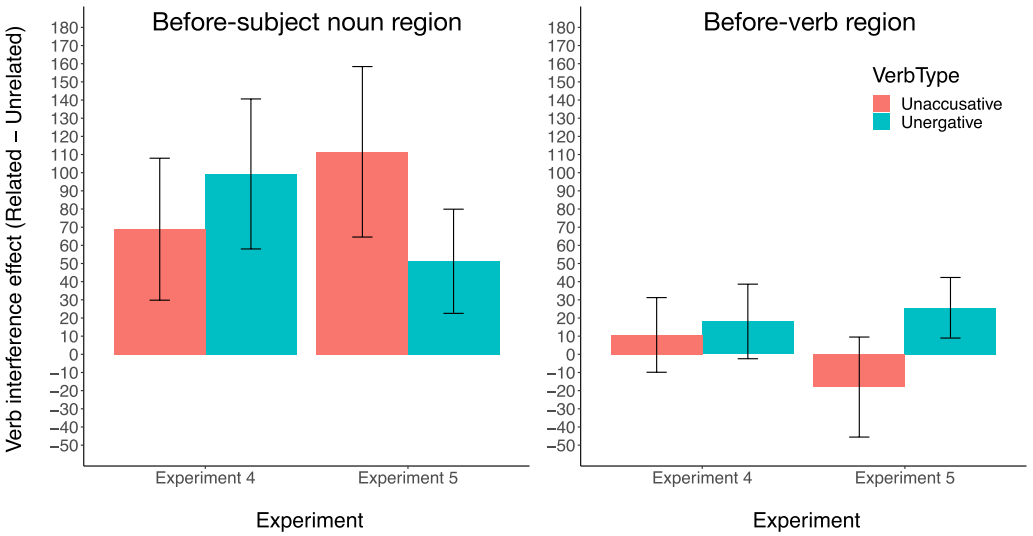

Fig. 22. The verb interference effects in Experiment 4 (without stop-signal task) and Experiment 5(with stop-signal task) in the before-subject noun region (left) and before-verb region (right). Error bars represent 95% confidence intervals.

grammatical encoding. The current results show that speakers are in fact capable of retrieving the verb before starting to articulate their internal arguments, without first retrieving linearly intervening words (here, the noun phrase inside subject-modifying adjuncts). Thus, it can be argued that speakers can prioritize that sentence-final verbs be retrieved before their sentence-initial internal arguments, over retrieving sentence-medial nouns which the head of internal arguments do not depend on.

Why are verbs needed to encode their internal arguments grammatically? There are several possible hypotheses (see Momma et al., 2018 for discussion), and we focus on five here. First, it is possible that speakers retrieve unaccusative verbs in advance to encode semantic dependencies between internal arguments and verbs. Under this account, the relational meaning of internal arguments (theme or patient arguments) is verb-dependent, but the relational meaning of external arguments (agent arguments) is verb-independent. Kratzer (p.4) captures the basic intuition behind this claim by, "Themes lack the conceptual independence of agents. Theme arguments seem to be tightly linked to their verbs. Agents are different. Actions seem to have agents independently of how we describe them." Thus, agent arguments are relatively independent of verbs, but theme or patient arguments are dependent on verbs, in the interface between syntax and semantics, or in the semantic representation.

If this representational difference in argument-verb relations is transparently reflected in sentence production, it may be possible to explain why speakers retrieve verbs before speaking their internal arguments but not before speaking their external arguments. Specifically, internal arguments require verb roots to obtain their relational meaning in a sentence, but external arguments do not need a verb root to obtain relational meaning (agent roles). Critically, semantic interference affects the retrieval of verb roots, because verb roots carry the verb meaning that the related (verb) distractor is similar to, and thus their retrieval is susceptible to semantic interference.

The second possibility is that speakers retrieve unaccusative verbs in advance to encode the syntactic dependency between internal arguments and verbs. Under this account, speakers need to retrieve verbs in advance to integrate the internal arguments into the sentence structure, because the internal structure of verb phrases is dependent on verbs' argument structures. In linguistics, it has been long observed that the phrase structure of a verb phrase depends on verbs' subcategorization (e.g., Chomsky, 1965). In some theories of syntax, the subject arguments of unaccusative verbs are underlyingly their objects and thus are a part of verb phrases (Perlmutter, 1968; Perlmutter, 1978). On this account, in order to syntactically encode the subjects of unaccusative verbs, it is necessary to retrieve verbs. This may be the cause of advance verb retrieval in sentences with unaccusative verbs.

The third possibility is that speakers retrieve unaccusative verbs in tandem with internal arguments because an internal argument and its verb, but not an external argument and its verb, form a semantic unit. Under this account, unaccusative verbs and their subject argument form integrated units at the level of semantic representation, and parts of a single semantic unit are planned in tandem. This account may be closely related to the notion of *semantic integratability* (Solomon & Pearlmutter, 2004), which is argued to influence the scope of planning in sentence production.

The fourth possibility is that speakers retrieve unaccusative verbs in tandem with internal arguments because internal arguments and verbs form a syntactic constituent. If this is the case, unaccusative verbs and their subject arguments form verb phrases in the underlying syntactic structure. Combined with the idea that speakers retrieve lemmas in advance within a first phrase (Smith & Wheeldon, 1999), it can be argued that speakers retrieve both the derived subject and verb in tandem, before starting to speak unaccusative sentences.

Finally, speakers may retrieve unaccusative verbs to choose an appropriate structure from multiple possible structural alternatives given a message. A message that is typically associated with sentences headed by unaccusative verbs can be expressed by multiple structures. For example, *the octopus is boiling* (a sentence headed by an unaccusative verb), *the octopus is being boiled* (a short passive

sentence) and *someone boils the octopus* (an active transitive sentence) arguably share the same message representation, so speakers must choose one structure among them. In contrast, a message that is typically associated with sentences headed by unergative verbs is associated with a smaller number of structures (e.g., there is no obvious alternative to *the octopus is swimming*). For example, most if not all the unaccusative verbs we used here can be passivized (see Appendix), while all unergative verbs we used cannot be passivized. This contrast may be responsible for the selective retrieval of unaccusative verbs; speakers might need to retrieve the verb to choose appropriate structures among multiple structures that are compatible with a message. Models of sentence production that are sensitive to the availability of structural alternatives (e.g., The Dual-path model as in Chang, 2002; Chang, Dell, & Bock, 2006) would naturally be compatible this possibility.

Thus, the current results can be explained in terms of semantic dependency, syntactic dependency, semantic constituency, or syntactic constituency, or the potential involvement of verbs in selecting a specific structure from multiple structural possibilities. Based on the current data, we cannot distinguish between these explanations. However, one way to empirically distinguish dependency-based and constituency-based accounts is to test whether speakers retrieve object arguments before producing verbs. Under the dependency-based hypothesis, speakers should not need to retrieve object nouns in advance, because verbs are not dependent on their object arguments. On the other hand, under the constituency-based hypothesis, it can be predicted that speakers need to retrieve the object noun before saying the verb. There is some evidence to suggest that speakers do not need to plan verbs when producing verb-phrase utterances. In an unpublished study, Momma et al. (2015) showed that semantic interference effects on object nouns in simple verb-phrase production (e.g., *feeding chickens*) did not appear in the onset latency measures, suggesting that speakers do not consistently retrieve the object noun before producing the verb. In contrast, they found reliable interference when nouns were produced in isolation. Given this result, we are currently partial to dependency-based accounts, though we remain agnostic about whether it is semantic dependency or syntactic dependency that triggers the advance verb planning.

In addition, if unaccusative verbs are needed to choose a structure given multiple structural possibilities, it is predicted that the unaccusative verbs that do not alternate (e.g., *arrive, appear, etc.*) should not have to be retrieved early. Thus, an experiment in which the alternating and non-alternating unaccusative verbs are contrasted can potentially rule out or support such an account. Also, as it currently stands, there is some empirical reason to disfavor this explanation. Ferreira (1996) showed that production is *easier* when multiple structure can be used to express a message, suggesting that the production of a particular structure is not based on competitive selection process (but see Hwang & Kaiser, 2014). If building sentence structures are not based on a competitive selection process, it is unclear why verbs are needed when there are multiple possible structures.

In sum, the precise reasons that speakers retrieve verbs before their internal arguments remain unspecified, but the above describes five hypotheses and some empirical ways to test them, and some empirical reasons to favor the dependency-based accounts. However, regardless of which account is correct, the current results suggest that the time course of speakers' planning of utterances reflects abstract relationships between arguments and their verbs, over and above surface word-order.

8.2. Is sentence production not incremental?

The current results may at first seem to be inconsistent with the widely accepted view that sentence production is incremental. But this depends on what it means for the sentence production system to be incremental. There are at least three distinct definitions of incrementality in the literature. First, at the most general level of description, incrementality refers to the idea that speakers can (but not must) interleave the planning and articulation of a single sentence (Ferreira & Swets, 2002; Levelt, 1989; Bock & Ferreira, 2013). This version of incrementality is not controversial in the literature and is not in conflict with the current results. In fact, the present findings support it by showing that speakers do not (need to) retrieve adjunct nouns and unergative verbs until immediately before they need to speak them.

Second, incrementality can refer to the idea that the retrieved words and planned structures are immediately integrated to the overall representation of a sentence (Momma & Phillips, 2018). This version of incrementality is also compatible with the current results: the unaccusative verbs that are retrieved in advance may be immediately integrated with the intermediate representation of a sentence without delay. Note that this notion of incrementality is parallel to the idea of incrementality in syntactic parsing (Momma & Phillips, 2018; Sturt & Lombardo, 2005; Demberg, Keller, & Koller, 2013). In parsing, incrementality refers to the property of the parser that immediately integrates the input representation to the overall representation of a sentence so that the syntactic representation is connected throughout a comprehension process.

Finally, incrementality can refer to the idea that planning and articulation of a particular element of an utterance synchronizes as much as possible (De Smedt, 1996; Van & Dietrich, 2003; Ferreira, 2000; Christiansen & Chater, 2016). The current results are not readily compatible with this version of incrementality, in the sense that the planning and articulation of unaccusative verbs was shown to be systematically de-synchronized. This version of incrementality is what gives rise to the sequential retrieval assumption, as we discussed in the introduction, and is primarily motivated by working memory considerations. A natural concern, then, is that advance unaccusative verb retrieval may not be cognitively economical. Thus, we next address how the current results relate to working memory in sentence production.

8.3. Working memory and sentence planning

As discussed in the introduction, a primary motivation for sequential word retrieval is to minimize memory cost during sentence production. Contrary to this memory consideration, the current results suggest that speakers retrieve unaccusative verbs that appear far downstream in the early stages of sentence planning. If speakers keep the retrieved verb in working memory until it can be

spoken, the process of producing unaccusative sentences (and also passive sentences, according to the previous results by Momma et al., 2015) may be costly. Also, if the current claims were to apply cross-linguistically, it is predicted that verb-final languages, all else being equal, are harder to speak than non-verb-final languages.

However, whether these predictions follow from the current claim depends on how memory cost is defined in sentence production. There are at least two reasons to postulate that retrieving verbs in advance might not be costly in terms of working memory. First, as long as the verbs that are retrieved in advance are immediately integrated into the overall structure of a sentence (in line with the second notion of incrementality discussed above), verbs become a part of a syntactic chunk, that is, a constituent. Just like chunking alleviates memory cost in other domains of cognition (Miller, 1956), it is possible that speakers may hold arbitrarily many words as long as they can form a single connected constituent structure. Under this view, a single syntactic chunk needs to be held in working memory throughout the course of producing a sentence, so sentence production is of course not cost free. But as a result of chunking, the cost of producing a sentence does not increase proportionally to the number of words. Thus, it is not clear if holding multiple words in memory is necessarily as costly in terms of working memory.

Second, verbs may not be a source of interference for other elements of a sentence. It is well-known in the working memory literature that the similarity between relevant items largely determines working memory cost. This is true for both capacity-based views (BBaddeley & Hitchaddeley & Hitch, 1974; Miller, 1956; Just & Carpenter, 1992) and interference-based views (Nairne, 1990). Generally speaking, when some item or items are held in working memory, maintenance and retrieval of other similar items can be harder. The view that similarity is a major factor affecting processing cost is widely accepted in psycholinguistics, in particular, sentence comprehension (Lewis & Vasishth, 2005; Van & Dietrich, 2003). Critically, it has been suggested that lemmas belonging to the different syntactic categories (e.g., nouns vs. verbs) may not interfere with each other in sentence production, even when they are conceptually similar to each other (Dell et al., 2008; Momma, Buffinton, Slevc, & Phillips, 2016). For example, Momma et al. (2016) showed that speakers are slower to say a word in a sentence (e.g., *singing*) when given a conceptually similar distractor word in immediate memory (*whistling*), but only when the conceptually similar distractor is perceived as belonging to the same syntactic category as the target. For example, when both the to-be-produced and the distractor words are nouns, as in *her skillful singing/whistling* or are verbs as in *she is skillfully singing/whistling*, they interfered with each other. Critically, when both the to-be-produced and the distractor words did not share the same syntactic category, they did not interfere with each other. Given this category specificity of retrieval interference, buffering verbs as nouns are processed (or vice versa) may not be cognitively costly. If similarity-based retrieval interference is a primary determinant of processing cost only when items match in syntactic category, non-sequential lexical retrieval processes may not be costly, as long as only one item of a certain category is held in memory. In other words, speakers may process sentences in "one-word-of-each-category-at-a-time" fashion. For example, retrieving a sentence-final verb first may not incur much processing cost on the retrieval and production of linearly preceding nouns. Thus, given what we know about memory cost, the (selective) advance verb planning mechanism we described here may not be particularly costly. Of course, this does not mean that advance verb planning is cost-free; indeed, the comparison between Experiment 4 and Experiment 5 suggests that adding a stop-signal task to the production task modulated advance unergative verb planning. However, holding verbs in memory while processing nouns may still be relatively less costly than, for example, holding nouns in memory while processing other nouns.

One prediction of this view is that relative clauses modifying subject nouns of unaccusative verbs (and also passive verbs) should be difficult to produce, because under the current view, unaccusative verbs and passive verbs need to be held in working memory while the verbs of subject-modifying relative clauses are retrieved. This prediction can be tested experimentally, and also using corpus studies. For example, relative clauses modifying subject nouns may be predicted to be rarer in unaccusative than in unergative sentences, under the linking hypothesis that more difficult structures are less likely to be produced. Future studies may test these predictions.

8.4. Flexibility in planning scope

From the early days of modern sentence production research, it has been noted that how much speakers plan at a time, that is, *scope of planning*, may be flexible (Levelt, 1989), and this insight has been verified experimentally (Ferreira & Swets, 2002; Griffin, 2001; Konopka, 2012; Wagner et al., 2010). Under this view, a natural question in the current context is whether speakers *must* retrieve verbs before starting to speak their internal arguments.

Phenomenologically, it is implausible that speakers *must* retrieve verbs before their internal arguments in all circumstances. For example, speakers can name an object that happens to come to mind for whatever reasons in a phrasal format (e.g., *the computer*) without having any idea about the continuation of the sentence. Speakers should have no problem determining the continuation of this sentence using unaccusative verbs (e.g., *the computer fell from the table*) *after* having uttered the noun phrase. Of course, this is a phenomenological observation that needs to be interpreted cautiously, but according to this intuition, we consider it unlikely that speakers *must* retrieve a verb before its internal arguments under all circumstances.

When, then, do speakers need to plan verbs before their internal arguments? One possibility is that speakers must retrieve a verb to speak a noun phrase realizing its relational meaning in a sentence, not as a fragment that can be repaired into an internal argument. When speakers produce a particular noun phrase as a fragment and later fix it into a full sentence, they need not retrieve verbs in advance to starting to speak their internal arguments. In contrast, when they speak internal arguments as internal arguments, speakers may need to retrieve verbs before their internal arguments. One empirical way to distinguish these two 'modes' of speaking would be to measure the fluency of speech. When speakers say a noun phrase as a fragment, not as an internal argument, there should be a measurable pause (or filler, like *um*) before being able to produce the verb. Thus, it is possible that speakers *must* retrieve verbs before starting to speak their internal arguments as internal arguments *in fluent speech*. But in either mode of speech, speakers may

need verbs to integrate internal argument to the overall structure of sentences before or after the production of internal arguments.

Relatedly, we consider it unlikely that speakers *must* retrieve unergative verbs and adjunct nouns on a just-in-time basis. Indeed, the cross-experiment analysis reported above showed that different speakers retrieve unergative verbs at different times. Speakers can certainly retrieve unergative verbs and adjunct nouns before starting a sentence if they so choose. Thus, our claim is *not* that just-in-time planning is architecturally necessary, but that speakers do not need to plan ahead words that do not participate in verb-argument dependency. When speakers retrieve unergative verbs and adjunct nouns is likely to be influenced by varieties of both linguistic and non-linguistic factors, including task-demands, sentence complexity, working memory capacity, among other factors. But in the current experiments, speakers showed consistent noun interference in the before-adjunct noun region, suggesting that they retrieve adjunct nouns on a just-in-time basis relatively consistently. Likewise, most speakers in the current experiments (around 80%) showed unergative verb interference in relatively late regions, suggesting that they retrieve unergative verbs on a just-in-time basis relatively consistently.

8.5. Cross-linguistic (in-) variability

Different languages have different surface grammatical properties, and it is eminently possible that when and why speakers retrieve verbs differ across languages. For example, it has been proposed that speakers rely less on verbs' argument structure in verb-final languages, like Korean (Hwang & Kaiser, 2014) and Japanese (Iwasaki, 2010, though see Momma et al. (2016) for alternative interpretations of these studies). Some learning-based models, such as the Dual-path model of sentence production, also implicate cross-linguistic differences in terms of what kind of information plays a role at each step in sentence planning (Chang, 2002; Chang et al., 2006).

However, the above-stated generalization that speakers retrieve verbs before their internal arguments is consistent with previous studies on the timing of verb planning, at least in English (Momma et al., 2015; Momma et al., 2018, cf. Hwang & Kaiser, 2014), German (Schriefers et al., 1998), Japanese (Momma et al., 2016) and Korean (Hwang & Kaiser, 2014). Although the conclusion reached in those studies was not framed in terms of the generalization stated above, the results of each study either show that (a) speakers do not seem to (consistently) retrieve verbs before the articulation onset of external (agent) arguments, (b) speakers seem to retrieve verbs in advance of the articulation onset of internal (theme or patient) arguments, or both. Of course, future cross-linguistic studies may find some results that implicate cross-linguistic differences in the timing of verb planning. For example, in languages that involve *quirky subjects*, such as Icelandic (Zaenen, Maling, & Thráinsson, 1985; Andrews, 1982; Schütze, 1993; Sigurðsson, 1992), the phonological form of the subject depends on the particular choice of the verb. Thus, in languages like Icelandic, it is more likely that speakers need to retrieve verbs before subjects, though the truly idiosyncratic case is argued to be never an external argument (e.g., Woolford, 2006). Also, in ergative-absolutive languages like Basque, the object of transitive verbs and the subject of intransitive verbs receive the same case. In those languages, it appears that at least the transitivity information of the verb needs to be already retrieved (though it is still not entirely clear if the verb root itself must be retrieved). Despite this potential variability in when speakers plan verbs, a theory that allows cross-linguistic differences is inherently harder to falsify than a theory that assumes cross-linguistic uniformity. Thus, until proven otherwise, we maintain the strong hypothesis that speakers retrieve verbs before their internal arguments, regardless of language.

8.6. Methodological remarks

The current study combines the word-by-word production time measurement (using an automatic forced alignment algorithm) with an interference paradigm to track the real-time cost associated with specific sub-processes in sentence production. Certainly, other methodologies, such as visual-world eye-tracking during speaking (e.g., Griffin & Bock, 2000 among others) also allows measurement of what speakers process in real-time, but the current method nicely complements them. The advantage of the current method is that it allows researchers to investigate the processing cost of sentence parts that do not directly correspond to an easily definable region in a picture, such as verbs (though see, e.g., Hwang & Kaiser, 2014 for an attempt). Most of the time these types of sentence parts encode abstract relational information, and arguably how speakers encode relational information is the critical but missing part of sentence production theories. It is possible to imagine extending the current paradigm to probe the timing of even more abstract processing, such as the processing of tense, empty categories, functional heads, and so forth, by carefully choosing the right kinds of distractors, or by inventing some manipulation that specifically affects relevant processes. This sort of investigation is not easily possible with existing methods, and thus the current method opens up new opportunities for investigating previously under-investigated aspects of how sentence planning unfolds over time in speaking. Of course, this methodology has some limitations; for example, as the current results suggest, the interference effect that occurs late in a sentence may not be reliably detected.

Also, in all experiments reported here, we pre-trained speakers to say a particular sentence for each picture. This pre-training procedure was necessary to ensure that speakers say the sentence we want them to say, and potentially even to elicit reliable semantic interference effects (Collina et al., 2013). It is possible that pre-training affects how speakers plan sentences. For example, pre-training might increase (or decrease) how many words speakers plan before starting to speak. Because the main focus of the current study is how speakers plan sentences, it is important to be cautious about generalizing the current results to everyday speaking. Relatedly, in the current experiments, the sentence structure, especially the subject-modifying prepositional phrase structure, was predicable (i.e., the subject-modifying prepositional phrase was always present in Experiment 1–3 and was always in the form of *X above/below Y*). It is possible that the predictable structures are planned less in advance, so the just-in-time noun interference effect was due to speakers being more willing to postpone the retrieval of nouns inside the predictable prepositional phrase (though it is

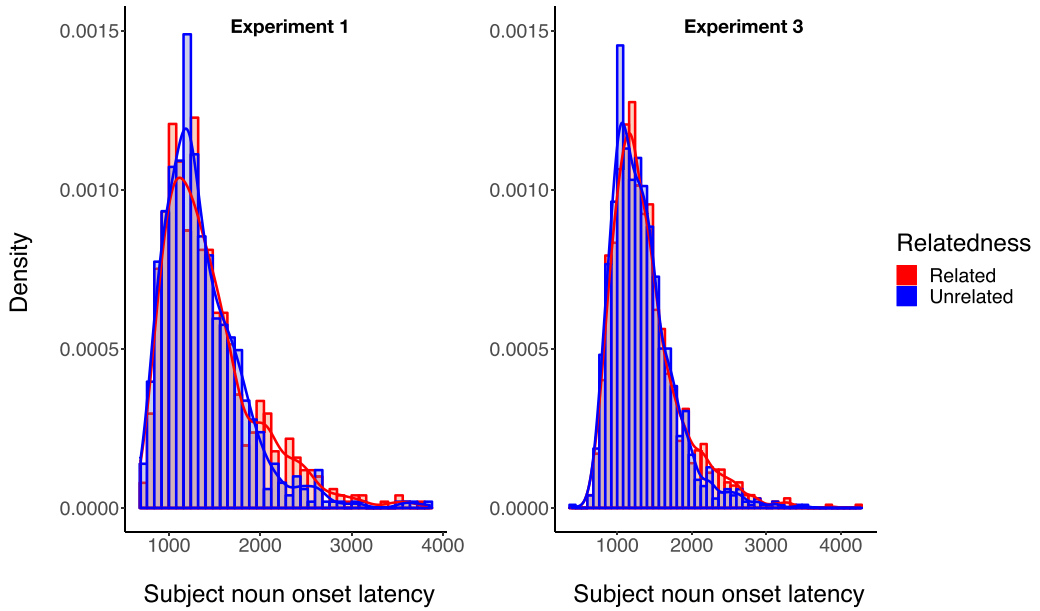

Fig. 23. The density plots visualizing the distribution of subject noun onset latency in the unaccusative verb distractor conditions by distractor relatedness in Experiment 1 and 3 (in −150 ms SOA condition).

equally possible that predictability encourages speakers to retrieve words more in advance). However, we should note that the contrast between unaccusative and unergative verbs we found in the current experiments still needs to be explained, even if the pre-training or predictability of structures affects how speakers plan sentences. There is no clear reason why pre-training or predictability affects sentences headed by unaccusative and unergative verbs differently. Nevertheless, converging results from naturalistic data (e.g., speech error data) or corpus analyses would be helpful to examine how the current results generalizes to everyday speaking.

Finally, in the current experiments, we analyzed our data using Inverse Gaussian mixed effects models, instead of, for example, more commonly adopted analyses methods like linear mixed effects analysis assuming log-normal distributions. This decision was based on previous observations (e.g., Momma et al., 2016) that semantic interference effects in picture-word interference tasks reside in the right tail of the production time distribution (i.e., in relatively slow trails). In fact, as can be seen in Fig. 23, the unaccusative interference effect we saw in the subject noun onset latency primarily reside in the right tail of the distribution, consistent with previous observations. In ideal situations, the analysis that can examine the skew of the distribution separately from central tendency like ex-Gaussian analyses (Luce, 1986; Heathcote, Popiel, & Mewhort, 1991) may be an even better tool for detecting the semantic interference effects and for understanding the cognitive underpinning of the effect. However, such analyses require more trials per participant per condition for the robust estimation of ex-Gaussian parameters (Ratcliff, 1979).

9. Conclusion

Since the seminal work by Garrett (1975), a major goal of sentence production research has been to understand how speakers translate conceptual representations into syntactic representations. Verbs' argument structures are a critical part of a semantic-syntax or conceptual-grammatical interface. Thus, they are likely to play a key role in this translation process. The current studies show that speakers retrieve words in sentences non-sequentially and that the order of lemma retrieval reflects verb-argument dependencies over and above surface word-order. Therefore, developing an adequate model of sentence production requires both incorporating theories of abstract argument structures and abandoning the default simplifying assumption that the time-course of lexical planning is transparently reflected in the surface word-order of a sentence.

Appendix A. Stimuli list

Table A1.

Table A1
The target sentences and distractor words used in Experiment 1–5.

Target sentence	Verb type	Rel. noun	Unrel. noun	Rel. verb	Unrel. verb
The ballerina above the ax/harp is shrinking	Unacc.	sword/violin	spoon/church	grow	float
The boy below the desk/lighthouse is floating	Unacc.	chair/windmill	violin/apple	drown	shrink
The chef above the dresser/windmill is drowning	Unacc.	table/lighthouse	hammer/chair	float	bounce
The clown below the violin/canon is growing	Unacc.	harp/gun	desk/lemon	shrink	shake
The cowboy above the sword/piano is falling	Unacc.	ax/guitar	drill/carrot	sink	boil
The dog above the hammer/apple is spinning	Unacc.	drill/lemon	dresser/lighthouse	trip	sink
The monkey above the carrot/knife is tripping	Unacc.	tomato/spoon	piano/castle	spin	melt
The octopus below the spoon/lemon is boiling	Unacc.	knife/apple	ax/cannon	melt	fall
The penguin below the drill/tomato is bouncing	Unacc.	hammer/carrot	sword/table	shake	drown
The pirate below the guitar/gun is sinking	Unacc.	piano/cannon	gun/guitar	fall	spin
The rabbit above the church/chair is shaking	Unacc.	castle/windmill	harp/windmill	bounce	grow
The snail below the castle/table is melting	Unacc.	church/dresser	knife/tomato	boil	trip
The ballerina above the ax/harp is running	Unerg.	sword/violin	spoon/church	swim	sneeze
The boy below the desk/lighthouse is yawning	Unerg.	chair/windmill	violin/apple	sleep	bark
The chef above the dresser/windmill is yelling	Unerg.	table/lighthouse	hammer/chair	bark	sleep
The clown below the violin/canon is walking	Unerg.	harp/gun	desk/lemon	crawl	cough
The cowboy above the sword/piano is winking	Unerg.	ax/guitar	drill/carrot	smile	crawl
The dog above the hammer/apple is barking	Unerg.	drill/lemon	dresser/lighthouse	yell	yawn
The monkey above the carrot/knife is sleeping	Unerg.	tomato/spoon	piano/castle	yawn	yell
The octopus below the spoon/lemon is swimming	Unerg.	knife/apple	ax/cannon	run	smile
The penguin below the drill/tomato is sneezing	Unerg.	hammer/carrot	sword/table	cough	run
The pirate below the guitar/gun is coughing	Unerg.	piano/cannon	gun/guitar	sneeze	walk
The rabbit above the church/chair is smiling	Unerg.	castle/windmill	harp/windmill	wink	swim
The snail below the castle/table is crawling	Unerg.	church/dresser	knife/tomato	walk	wink

References

Alexiadou, A., Anagnostopoulou, E., & Everaert, M. (2004). *The unaccusativity puzzle: Explorations of the syntax-lexicon interface*, *vol. 5*. Oxford University Press on Demand.

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.

Andrews, A. (1982). The representation of case in modern icelandic. *The Mental Representation of Grammatical Relations, 361*, 427–503.

Baddeley, A. D. (2010). Working memory. *Current Biology, 20*(4), R136–R140.

Baddeley, A. D., & Hitch, G. (1974). Working memory. *Psychology of learning and motivation: vol. 8*, (pp. 47–89). Elsevier.

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.

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, K., & Ferreira, V.S. (2013). Syntactically speaking. In *The oxford handbook of language production*.

Bock, K., & Levelt, W. (2002). Language production. *Psycholinguistics: Critical concepts in psychology, 5*, 405.

Bock, K., & Levelt, W. J. (1994). Language production: Grammatical encoding. *Handbook of psycholinguistics* (pp. 945–984). Academic Press.

Brysbaert, M., & New, B. (2009). Moving beyond kučera and francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41*(4), 977–990.

Chang, F. (2002). Symbolically speaking: A connectionist model of sentence production. *Cognitive Science, 26*(5), 609–651.

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review, 113*(2), 234.

Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT.

Chomsky, N. (1986). *Barriers*, *vol. 13*. MIT press.

Chomsky, N. (1986). *Lectures on government and binding: The pisa lectures*, *vol. 9*. Walter de Gruyter.

Chomsky, N., & Keyser, S. J. (1982). *Some concepts and consequences of the theory of government and binding*. MIT press.

Christiansen, M. H., & Chater, N. (2016). The now-or-never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences,* 39.

Christianson, K., & Ferreira, F. (2005). Conceptual accessibility and sentence production in a free word order language (odawa). *Cognition, 98*(2), 105–135.

Collina, S., Tabossi, P., & De Simone, F. (2013). Word production and the picture-word interference paradigm: The role of learning. *Journal of Psycholinguistic Research, 42*(5), 461–473.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93*(3), 283–321.

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.

Demberg, V., Keller, F., & Koller, A. (2013). Incremental, predictive parsing with psycholinguistically motivated tree-adjoining grammar. *Computational Linguistics, 39*(4), 1025–1066.

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, F., & Swets, B. (2002). How incremental is language production? Evidence from the production of utterances requiring the computation of arithmetic sums. *Journal of Memory and Language, 46*(1), 57–84.

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., & Dell, G. S. (2000). Effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology, 40*(4), 296–340.

Fox Tree, J. E., & Clark, H. H. (1997). Pronouncing the as thee to signal problems in speaking. *Cognition, 62*(2), 151–167.
Garrett, M. (1975). The analysis of sentence production. In H. Bower (Vol. Ed.), *The psychology of learning and motivation: vol. 9*, (pp. 133–177). New York, NY: Academic Press.
Gordon, J. K., & Dell, G. S. (2003). Learning to divide the labor: An account of deficits in light and heavy verb production. *Cognitive Science, 27*(1), 1–40.
Griffin, Z. M. (2001). Gaze durations during speech reflect word selection and phonological encoding. *Cognition, 82*(1), B1–B14.
Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science, 11*(4), 274–279.
Grimshaw, J. (1990). *Argument structure*. MIT Press.
Haegeman, L. (1991). *Introduction to government and binding theory, vol. 2*. Blackwell Oxford.
Heathcote, A., Popiel, S. J., & Mewhort, D. (1991). Analysis of response time distributions: An example using the stroop task. *Psychological Bulletin, 109*(2), 340.
Hwang, H., & Kaiser, E. (2014). Having a syntactic choice is not always better: the effects of syntactic flexibility on korean production. *Language, Cognition and Neuroscience, 29*(9), 1115–1131.
Hwang, H., & Kaiser, E. (2014). The role of the verb in grammatical function assignment in english and korean. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(5), 1363.
Indefrey, P., & Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition, 92*(1–2), 101–144.
Iwasaki, N. (2010). Incremental sentence production: Observations from elicited speech errors in japanese. *Processing and producing head-final structures* (pp. 131–151). Springer.
Jescheniak, J. D., & Levelt, W. J. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(4), 824.
Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*(1), 122.
Kaplan, R. M., & Bresnan, J. (1982). Lexical-functional grammar: A formal system for grammatical representation. *Formal Issues in Lexical-Functional Grammar, 47*, 29–130.
Kempen, G., & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science, 11*(2), 201–258.
Kempen, G., & Huijbers, P. (1983). The lexicalization process in sentence production and naming: Indirect election of words. *Cognition, 14*(2), 185–209.
Kittredge, A. K., Dell, G. S., Verkuilen, J., & Schwartz, M. F. (2008). Where is the effect of frequency in word production? Insights from aphasic picture-naming errors. *Cognitive Neuropsychology, 25*(4), 463–492.
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. *Phrase structure and the lexicon* (pp. 109–137). Springer.
Kratzer, A. (2003). *The event argument and the semantics of verbs*. Ms.Amherst: University of Massachusetts.
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.
Lappin, J. S., & Eriksen, C. W. (1966). Use of a delayed signal to stop a visual reaction-time response. *Journal of Experimental Psychology, 72*(6), 805.
Levelt, W.J. (1989). Speaking: From intention to articulation. A Bradford book; *ACL-MIT Press series in natural-language processing*.
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.
Levin, B., & Hovav, M. R. (1995). *Unaccusativity: At the syntax-lexical semantics interface, vol. 26*. MIT press.
Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science, 29*(3), 375–419.
Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology, 6*, 1171.
Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review, 91*(3), 295.
Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization, vol. 8*. Oxford University Press on Demand.
Lupker, S. J. (1979). The semantic nature of response competition in the picture-word interference task. *Memory & Cognition, 7*(6), 485–495.
MacKay, D. G. (1982). The problems of flexibility, fluency, and speed–accuracy trade-off in skilled behavior. *Psychological Review, 89*(5), 483.
Marantz, A. (1981). *On the nature of grammatical relations*. Unpublished doctoral dissertationMassachusetts Institute of Technology.
Martin, R. C., & Freedman, M. L. (2001). Short-term retention of lexical-semantic representations: Implications for speech production. *Memory, 9*(4–6), 261–280.
Matzke, D., Verbruggen, F., & Logan, G. D. (2018). The stop-signal paradigm. Stevens' Handbook of Experimental Psychology and Cognitive. *Neuroscience, 5*, 1–45.
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.
Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*(2), 81.
Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex frontal lobe tasks: A latent variable analysis. *Cognitive Psychology, 41*(1), 49–100.
Momma, S., Buffinton, J., Slevc, L.R., & Phillips, C. (2016). *Syntactic category constrains lexical competition in speaking*.
Momma, S., & Phillips, C. (2018). The relationship between parsing and generation. *Annual Review of Linguistics, 4*.
Momma, S., Slevc, L., & Phillips, C. (2015). The timing of verb planning in active and passive sentence production. *Poster presented at the 28th annual cuny conference on human sentence processing, Los Angeles, CA, March 19–21*.
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.
Momma, S., Slevc, L. R., & Phillips, C. (2018). Unaccusativity in sentence production. *Linguistic Inquiry, 49*(1), 181–194.
Nairne, J. S. (1990). A feature model of immediate memory. *Memory & Cognition, 18*(3), 251–269.
Oldfield, R. C., & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology, 17*(4), 273–281.
Perlmutter, D. M. (1968). *Deep and surface structure constraints in syntax*. Unpublished doctoral dissertationMassachusetts Institute of Technology.
Perlmutter, D.M. (1978). Impersonal passives and the unaccusative hypothesis. In *Annual meeting of the berkeley linguistics society* (vol. 4, pp. 157–190).
Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin, 86*(3), 446–461.
Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition, 42*(1–3), 107–142.
Sag, I. A., & Pollard, C. J. (1987). *Head-driven phrase structure grammar*. Morgan Kaufmann.
Schnur, T. T. (2011). Phonological planning during sentence production: Beyond the verb. *Frontiers in Psychology, 2*, 319.
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.
Schütze, C. T. (1993). Towards a minimalist account of quirky case and licensing in icelandic. *MIT Working Papers in Linguistics, 19*, 321–375.
Sigurðsson, H.Á. (1992). Verbal syntax and case in icelandic: In *A comparative gb approach*. Citeseer.
Slevc, L. R. (2011). Saying what's on your mind: Working memory effects on sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(6), 1503.
Smith, M., & Wheeldon, L. (1999). High level processing scope in spoken sentence production. *Cognition, 73*(3), 205–246.
Solomon, E. S., & Pearlmutter, N. J. (2004). Semantic integration and syntactic planning in language production. *Cognitive Psychology, 49*(1), 1–46.
Steedman, M. (2000). *The syntactic process*. MIT press.
Sturt, P., & Lombardo, V. (2005). Processing coordinated structures: Incrementality and connectedness. *Cognitive Science, 29*(2), 291–305.
Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., Bates, E., et al. (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory and Language, 51*(2), 247–250.
Van Nice, K. Y., & Dietrich, R. (2003). Task sensitivity of animacy effects: Evidence from german picture descriptions. *Linguistics, 41*(5; ISSU 387), 825–850.

Verbruggen, F., & Logan, G. D. (2008). Response inhibition in the stop-signal paradigm. *Trends in Cognitive Sciences, 12*(11), 418–424.
Vigliocco, G., Vinson, D. P., Lewis, W., & Garrett, M. F. (2004). Representing the meanings of object and action words: The featural and unitary semantic space hypothesis. *Cognitive Psychology, 48*(4), 422–488.
Vince, M. A. (1948). The intermittency of control movements and the psychological refractory period. *British Journal of Psychology, 38*(3), 149.
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.
Woolford, E. (2006). Lexical case, inherent case, and argument structure. *Linguistic Inquiry, 37*(1), 111–130.
Yamashita, H., & Chang, F. (2001). long before short preference in the production of a head-final language. *Cognition, 81*(2), B45–B55.
Yuan, J., & Liberman, M. (2008). Speaker identification on the scotus corpus. *Journal of the Acoustical Society of America, 123*(5), 3878.
Zaenen, A., Maling, J., & Thráinsson, H. (1985). Case and grammatical functions: The icelandic passive. *Natural Language & Linguistic Theory, 3*(4), 441–483.