



ELSEVIER

Contents lists available at [ScienceDirect](#)

Cognitive Psychology

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



Language knowledge and event knowledge in language use



Jon A. Willits^{a,*}, Michael S. Amato^b, Maryellen C. MacDonald^c

^a Indiana University, United States

^b Clearway Minnesota, United States

^c University of Wisconsin Madison, United States

ARTICLE INFO

Article history:

Accepted 6 February 2015

Keywords:

Language comprehension
Language production
Semantic memory
Word meaning
Event representation

ABSTRACT

This paper examines how semantic knowledge is used in language comprehension and in making judgments about events in the world. We contrast knowledge gleaned from prior language experience (“language knowledge”) and knowledge coming from prior experience with the world (“world knowledge”). In two corpus analyses, we show that previous research linking verb aspect and event representations have confounded language and world knowledge. Then, using carefully chosen stimuli that remove this confound, we performed four experiments that manipulated the degree to which language knowledge or world knowledge should be salient and relevant to performing a task, finding in each case that participants use the type of knowledge most appropriate to the task. These results provide evidence for a highly context-sensitive and interactionist perspective on how semantic knowledge is represented and used during language processing.

© 2015 Published by Elsevier Inc.

1. Introduction

One of the primary purposes of language is to communicate about entities and events in the world. Therefore, language use necessarily involves integration of knowledge about linguistic forms and

* Corresponding author at: Department of Psychological and Brain Sciences, Indiana University, 1102 E. 10th St., Bloomington, IN 47406, United States.

E-mail address: jwillits@indiana.edu (J.A. Willits).

knowledge of those forms' real world referents. For example, if a speaker is telling a listener about a squirrel, the listener uses both world knowledge and language knowledge to comprehend the speaker's utterances. Examples of relevant world knowledge include prior encounters with squirrels and direct experience with their characteristic behaviors. An example of language knowledge is phonotactic knowledge about English, such as the fact that the phoneme sequence /skw/ is a rare but permissible sequence in English, and that the sequence is most likely to be at word onset, as in the word *squirrel*. In these examples, the world knowledge (observations of an animal) and language knowledge (experience with the phoneme sequences in the word that names the animal in English) are very easily distinguished. They come from different experiences (have a different ontogenesis) and they are used for different tasks, such as reasoning about squirrels vs. recognizing the word *squirrel* in the speech stream.

In other cases, however, world and language knowledge become easily blurred. For example, perhaps the comprehender brings to bear some information about squirrels that came not from direct experience but from reading or hearing something about squirrels, such as having read, "Squirrels bury nuts in the yard." This information has elements of both world and language knowledge; comprehenders of this sentence receive information about squirrel behavior in the world, but the experience of reading the sentence also provides language knowledge, such as about the co-occurrence of words, as in the trigram *squirrels bury nuts*. This sequential word co-occurrence information, like sequential phoneme information (phonotactics), affects patterns of reading and language comprehension (McDonald & Shillcock, 2003).

This article is aimed at elucidating the relationship between the *ontogenesis* of knowledge and the extent to which world-derived and language-derived knowledge are brought to bear under different task demands. As we've just noted, there are many varieties of world knowledge and language knowledge, but our explorations will focus on particular types. On the world side, we investigate probabilities associated with people's knowledge of events, such as the probability that squirrels bury things, or the probability that a burying event takes place in a yard. On the language side, the knowledge we investigate is knowledge of word co-occurrences, such as the co-occurrence of *squirrel* and *bury* or *bury* and *yard*. These two instantiations do not exhaust either the world or language knowledge that people possess, but they can be aligned in a way that permits useful comparisons. In the next section, we argue for the importance of this world/language knowledge distinction in current theories of language comprehension. We then present four experiments investigating how the balance between us of language (word co-occurrence) knowledge and world (event) knowledge changes as a function of varying task demands.

1.1. The growing importance of language knowledge in language comprehension

While researchers have always identified clear cases of language knowledge such as phonotactics, and similarly clear cases of world knowledge, until recently subtle language/world distinctions with respect to semantic knowledge have not been of primary concern to accounts of language comprehension. Modular accounts of language comprehension that were developed in the 1970s were notable in distinguishing stages of processing (Frazier & Fodor, 1978; Swinney, 1979), but these stages did not cleave cleanly along language/world knowledge boundaries. For example, in Frazier's (1987; Frazier & Fodor, 1978) account of sentence interpretation, language-specific syntactic information guided initial parsing via Minimal Attachment and other parsing principles (language knowledge), and a later stage integrated this parse with semantic knowledge, without regard to whether semantic knowledge was world- or language-derived. More recently, constraint-based accounts of language comprehension have argued against distinct stages of processing and emphasized how the rapid use of complex probabilistic semantic information shapes sentence comprehension (MacDonald, Pearlmutter, & Seidenberg, 1994; McRae, Spivey-Knowlton, & Tanenhaus, 1998), again without clear claims about the ontogenesis of the information. For example, constraint-based approaches to language processing that have measured the effect of semantic plausibility on comprehension (as in studies that ask questions such as "How likely is it for a cop to arrest someone?") (Ferretti, McRae, & Hatherell, 2001) appear to be studying the effect of world knowledge on comprehension (e.g., the likelihood of a cop performing an arresting event in the world). However, these real-world probabilities are also reflected in linguistic

structure, such as the frequency of the noun *cop* occurring with or taking an agent role for the verb *arrest*, and the similarity of the contexts in which these words jointly occur (Jones, Kintsch, & Mewhort, 2006; Willits, Duran, D'Mello, & Olney, 2007). Similarly, MacDonald (1993) argued that language knowledge, specifically the frequency with which a noun served as the head of a noun phrase (e.g., *warehouse* in *the large warehouse*) vs. as the modifier of another head noun (*the warehouse manager*), affected online noun phrase interpretation and lexical category ambiguity resolution. Again, the language and world knowledge are confounded here: warehouses, being highly interconnected entities, have managers, policies, locations and relations to many other concepts that could be at the root of the language statistics that MacDonald observed. Thus, correlations between language and world knowledge make it difficult to identify both how information came to be acquired and the types of information that are routinely used in language comprehension.

Still more recently, an important shift in comprehension research has emphasized the role of knowledge about language distributions in comprehension processes such as reading. Some of these studies contrast properties of the grammar to event knowledge, without particular commitments to the extent to language or world statistics contribute to this knowledge (Patson & Warren, 2014). Other work has explicitly addressed statistical patterns in language input. Several studies have confirmed that high frequency word collocations (i.e. phrases in which the words have high transition probabilities) have different patterns of comprehension (Arnon & Snider, 2010; McDonald & Shillcock, 2003; Reali & Christiansen, 2007) or production (Wasow, 1997) than ones with lower collocation frequencies. Of course collocations necessarily reflect what people talk about and thus must indirectly reflect world knowledge, but it is difficult to dismiss the central role of language knowledge in these cases. For example, the Reali and Christiansen materials differed only in placement of pronouns, and the high vs. low collocation items in Arnon & Snider were all fairly semantically neutral and plausible pairs of phrases such as *I want to sit* vs. *I want to say*. Perhaps the clearest examples of this perspective are information theoretic approaches to reading processes, in which reading times are predicted to vary with words' predictability in context, owing to prior reading experience (Amato & MacDonald, 2010; Hale, 2001; Levy, 2008; Smith & Levy, 2013). Together, results from these studies suggest that comprehension is influenced by knowledge of transition probabilities across words and other language-specific statistics, gleaned from prior experience with language, above and beyond experience with nonlinguistic events in the world.

1.2. Language statistics and semantic representations

Beyond its role in language comprehension, language knowledge also affects the representation of semantic memory, which is commonly taken to encompass both linguistic and nonlinguistic aspects of conceptual representations. Statistical regularities in language have long been implicated as a key source of semantic knowledge. Dating back at least to Bloomfield (1933), Harris (1957), and Firth (1957), distributional tests were critical to early theories of phonological and grammatical categories. According to these theories, what makes a word like *car* a noun is that it follows articles like *the*, and not auxiliaries like *is*. A similar point has been made in behavioral studies of the syntactic bootstrapping of a word's meaning from its distribution of occurrence in particular morphological, lexical, and syntactic frames (Fisher, Gleitman, & Gleitman, 1991; Landau & Gleitman, 1985; Naigles, 1990; Lany & Saffran, 2010). Recent analyses of spoken and written languages have demonstrated that the distributional similarity of words, such as the number of contexts the two words share in common, is sufficient for clustering words into categories that are quite similar to what we think of as grammatical categories (Kiss, 1973; Maratsos & Chalkley, 1980; Mintz, 2003; Monaghan, Chater, & Christiansen, 2005; Redington Chater, & Finch, 1998). The distributional similarity of two words also has been shown to predict results in many experimental paradigms that are central to studies of semantic representations within cognitive science, including semantic priming (Jones, Mewhort, & Kintsch, 2006; Lund, Burgess, & Audet, 1996); semantic categorization (Riordan & Jones, 2010); noun-verb semantic relations (Hare, Jones, Thomson, Kelly, & McRae, 2009; Willits, Duran, D'Mello, & Olney, 2007), the development of vocabulary knowledge (Landauer & Dumais, 1997) and the development of hierarchically structured semantic representations (Willits & Jones, in review). Word statistics even predict the nature of semantic relationships that are typically thought of as embodied and therefore

drawing on real world experiences (Louwerse, 2008). Together, these results suggest that at least some semantic relationships may be computed using language knowledge.

To summarize these points, we have two related observations: *First*, studies of language comprehension have increasingly pointed to knowledge of language statistics as being crucial in accounts of comprehension processes. *Second*, studies of semantic memory and conceptual representations have also increasingly suggested that key aspects of semantic representations that have been considered as owing to experience with entities and events in the world might instead owe at least in part to experience with language. These two observations, together with evidence that world and language knowledge are correlated, could make it seem difficult to distinguish world and language knowledge in language use and semantic representations.

However, disentangling these two knowledge types is made possible by the fact that the correlation between the two types of information is not perfect. In fact, the structure of language input often significantly deviates from the world it is being used to describe, with important consequences for comprehension processes and for how semantic knowledge is represented. For example, when people talk about events in the world, their language routinely leaves out some details that are assumed, superfluous, or easily observed, as described by Grice (1975). Thus the utterance “I went bowling with Jerry last night” felicitously omits redundant information, such as the fact that the event (probably) involved bowling balls, pins, and shoes. Likewise, because of language’s communicative purpose, it is often used to note unusual or atypical facts: a speaker is much more likely to say “I stirred the soup with my finger” rather than “I stirred the soup with my spoon”, even though the latter is the more probable event in the world. Similarly, in contrasts like “I want to sit” vs. “I want to say” in Arnon and Snider (2010), comprehension time differences reflect word co-occurrences and thus implicate events that are more likely to be talked about, not necessarily events that are more likely to occur. Thus, comprehension processes that develop expectations for upcoming input will make very different predictions depending on whether the statistics of language or of the world are being used to generate those predictions. Similarly, conceptual knowledge or a semantic memory derived exclusively from linguistic structure would end up dramatically misrepresenting the likelihood of events in the world.

In the present work, we use disparities between events in the world and language statistics to take further steps toward understanding how input from language and experience with the world contribute to language comprehension and broader semantic tasks (Ferretti, Kutas, & McRae, 2007). In particular, we investigate an interesting set of findings on the processing of verbs and their thematic relations (e.g., nouns) within sentences. We then present two corpus studies and four behavioral experiments that address the question of the differential use of language knowledge and world knowledge across different task demands.

1.3. Verbs and the activation of event knowledge

In contrast to work discussed above emphasizing the role of language input in reading and comprehension, some theories of language comprehension emphasize the role of world knowledge in comprehension processes. These approaches hold that understanding a sentence involves the creation of a situation model or mental simulation (Barsalou, 1999; Johnson-Laird, 1983; Madden & Zwaan, 2003; Magliano & Schleich, 2000; van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998) or the activation of an event schema (Ferretti et al., 2001; Rumelhart & Levin, 1975). According to these accounts, as comprehenders encounter each word in an input stream, they gradually update their situation model or event schema based on how the new word’s referent interacts with the current model or schema. On this view, the semantic component in sentence comprehension involves learning the mapping between words or constructions in the linguistic input and fully realized event representations.

As an example of this perspective on language processing, Ferretti et al. (2007) investigated how a verb’s aspect affects the activation of event schema to which the verb refers. A verb’s aspect and tense together identify the time of a described event and its extension, (i.e., whether the event is extended in duration or not). Reichenbach (1947) notes that we can understand various tense/aspect combinations with respect to three time points: the time of the speaking, the reference time (the time of the broader topic, such as a story someone is telling), and the time of the specific event

denoted by the verb. For example, in “Mary was tired because the rooster had crowed very early this morning”, the verb form “had crowed” is in the past perfect form—it refers to a completed event (rooster crowed) that occurred farther back in time than the reference point of the story (that Mary was tired). By contrast, the past progressive aspect denotes an event in the past that was extended in duration: Mary was tired because the rooster was crowing very early this morning. Here the crowing event extends for some time but is again taking place before the reference point (the description of Mary’s fatigue). Thus the past perfect and the past progressive both refer to events prior to the reference point (and prior to the time at which the sentence is uttered), but they differ in aspect—the duration of that past event.

Ferretti et al. (2007) investigated effects of aspect on noun recognition in past perfect and past progressive verb phrases.¹ They found that past progressive forms (such as *was cooking*) facilitated responses to location nouns (like *kitchen*) in a word naming task, compared to an unrelated prime condition (e.g. *was cheering*). In contrast, they found that past perfect forms (such as *had cooked*) did not facilitate responses to locations. In both cases the basic semantic relationship between the prime and target was equivalent (TO COOK → KITCHEN), but priming of locations occurred for semantically related verb phrases only in the past progressive (*was X-ing*) form, not in the past perfect (*had X-ed*) form. Ferretti et al. argued that these aspect effects arise due to the morphosyntactic form differentially activating the semantic features or associated object concepts that are “consistent with the temporal reference of the different aspect categories” (pg. 183). Thus, past progressive verbs (e.g. *was X-ing*) affect the comprehender’s event schema or situation model such that certain nouns (like locations) are more active than they would be had the comprehender heard a past perfect verb (e.g. *had X-ed*).

Thus on this view, language comprehension draws on event knowledge in real time. The facilitated processing of locations following verbs describing ongoing actions – such as *was cooking* – stems from a mapping from the syntactic form of the verb to a subset of world knowledge about cooking events that are associated with that syntactic form: the verb activates an “ongoing-event” schema, which in turn activates associated location concepts, which then activate their lexical labels, resulting in facilitated processing of location words if they are subsequently encountered in the sentence. Ferretti et al. argued that the function of this process is to activate the event participants that are likely to be the most salient and relevant, as locations are in the case of ongoing events.

An alternative view is that Ferretti et al.’s results stem at least in part from knowledge of linguistic collocations between verb forms and location phrases. If a verb referring to an ongoing event facilitates processing of a location more than a verb form referring to a completed event, this result could be due to knowledge of language *usage* rather than world knowledge. Specifically, both ongoing and completed events necessarily took place in some location in the world, but the probability of *mentioning* a location may be higher when referring to ongoing actions than to completed events. Thus, the difference between the two accounts turns on both the ontogeny of the information and the functional role that the information is playing during language processing. Under the event knowledge account, the critical difference in responses to “*was X-ing*” verbs and “*had X-ed*” verbs is due to differences in representations of ongoing and completed events, and the extent to which locations are salient, associated, or structurally represented as part of one’s representations of the actual event, as it happened in the world. In the language knowledge account, the critical differences is due to differences in the representations of the verbs, or expectations generated about these verbs, based on prior language experience.

A final important note is that a distinction between an account ascribing behavioral effects to application of world knowledge vs. language knowledge is orthogonal to the question of the specificity of that knowledge. The knowledge about the effect of a verb’s aspect could be lexically independent – encountering *any* verb describing an ongoing or completed action (e.g. the *was X-ing* or the *had X-ed* frame) necessarily activates a world knowledge schema about events and/or changes the likelihood of location mentions in the language. Alternatively, under both accounts the knowledge could be lexically/event specific, bound to the representations of particular verb – encountering “*was cooking*” may

¹ Ferretti et al. referred to the distinction as imperfect/perfect, which refers to a combination of tense (whether the event happened in the past) and aspect (whether the event had extended duration). To avoid potential confusion, we will use the more standard linguistic terms past progressive (for *was X-ing*) and past perfect (for *had X-ed*).

activate locations like *kitchen* (because of a high degree of association, either in language or in the world), whereas “*was cleaning*” may *not* activate locations (perhaps because cleaning happens in or is talked about in reference to many locations).

In its language-knowledge instantiation, the lexically independent view is similar to other arguments in the sentence processing literature for what might be called “construction-wide” probabilistic constraints on interpretation, such as Spivey-Knowlton and Sedivy’s (1995) claim that definite and indefinite noun phrases had different probabilities of being subsequently modified with a prepositional phrase (e.g. “a/the pizza with pepperoni”), independent of the identity of the noun (*pizza*, in this case). However, a language knowledge account also admits a lexically-specific position, such that individual past progressive (i.e. “was X-ing”) verb forms tend to be more associated with certain location words than particular past perfect (i.e. “had X-ed”) verb forms. On this view, linguistic forms may still map to world knowledge, but the processes driving the speeded recognition of locations in Ferretti et al.’s study, and potentially in much of language comprehension, employ knowledge of the collocations of particular words in the language (Arnon & Snider, 2010), including inflected verb-location pairs. And of course lexically-specific and more construction-general hypotheses are not mutually exclusive, and interactions between them are attested in the literature. These results are often described as frequency-by-regularity interactions, where a lexically-specific effect and a construction-wide effect have different influences depending on the frequency of the elements involved (Juliano & Tanenhaus, 1993; Pearlmutter & MacDonald, 1995).

This discussion suggests that Ferretti et al.’s (2007) results, which demonstrated a link between past progressive verbs and locations, admits four different hypothetical explanations arising as interactions of two distinct factors. The first factor is whether the represented knowledge leading to the difference is a difference in language knowledge or world knowledge (e.g. knowledge about the words and their sentence environments, or knowledge about their referents in the world, respectively). The second factor is whether that knowledge is highly abstracted and rule-like, or verb/event specific. In other words, are people representing knowledge of an abstract relation between past progressive verbs (or ongoing events) and locations? Or are they representing these relations individually and possibly differently for each verb/event → location pair? At the heart of these questions are issues concerning the nature of the information that people extract from their previous experiences and deploy in interpretation of new input.

In the following corpus analyses and behavioral experiments, we used Ferretti et al.’s effects as a starting point for contrasting the role of language and event knowledge, as well as exploring the lexical dependence of these effects. In Corpus Study 1, we examine the extent to which the relationship between verb aspect and noun arguments observed by Ferretti et al. (2007) is reflected in the distributional patterns in the English language, as a first step to investigating the extent to which the effects ascribed by Ferretti et al. to event knowledge may also be contained in word-specific co-occurrences in English. Corpus Study 2 extends the investigation beyond the particular noun–verb pairings used by Ferretti et al., yielding a more construction-wide characterization of verb aspect and location arguments in English. We then present four behavioral experiments investigating the relative weight of language and world knowledge in processing verb-noun relationships. More specifically, we use the corpus results to select stimulus items for which the distributional patterns in the language conflict with the typical event knowledge patterns in the world, as identified by Ferretti and others, and we manipulate task demands that may tend to favor the use of world vs. language knowledge.

2. Corpus Study 1

The goal of Corpus Study 1 was to test the potential of one language knowledge-based explanation of the effects in Ferretti et al. (2007), specifically an account of those effects based on people’s knowledge of the co-occurrence probability of verb phrases and location arguments. In other words, we ask whether the facilitation effects in Ferretti et al.’s experiment can be explained by which verb phrases and locations co-occur frequently in language. This operationalization of language knowledge as co-occurrence is not an endorsement of language knowledge being limited to word co-occurrences. It is, however, a straightforward language analogue to the event (world) knowledge that Ferretti et al.

posit, and if co-occurrences offer an adequate account of the behavioral effects, then it may be possible to argue that (in this case) more complex representations are not necessary.

Ferretti et al.'s Experiment 1 used a priming paradigm, in which subjects were presented with a verb phrase as a prime (such as *was cooking*; note that we are using “verb phrase” here to refer only to the past progressive and past perfect verb forms such as *was cooking* and *had cooked* and not to other arguments/adjuncts of a verb such as its direct object or location). Participants then named a target word that appeared on a screen. For each target location, the experimental materials crossed verb aspect (either past progressive: *was cooking*; or past perfect: *had cooked*) with verb-location semantic relatedness (e.g. *was cooking* → *kitchen* vs. *was applauding* → *kitchen*). Ferretti et al. found an interaction, such that the semantically related primes facilitated responses in the past progressive condition (e.g. *was cooking* primed *kitchen* relative to *was applauding*) but not the past perfect condition (e.g. *had cooked* did not prime *kitchen* relative to *had applauded*). They argued that this response was due to differential activation of “ongoing” and “completed” event schemas, respectively. They rejected a role for language statistics because (1) aspect is traditionally not a part of spreading activation models (Collins & Loftus, 1975; Katz & Fodor, 1963; though see Gentner, 1981), and (2) their aspect-relatedness interaction could not be explained by differences in word association strength, as measured in large databases of word association norms (such as Nelson, McEvoy, & Schreiber, 1999). However, there is no reason why spreading activation models must exclude aspect information, nor is there any reason for language knowledge to be limited to simplistic spreading activation models. Moreover, large databases of word association norms severely underestimate the degree and types of relations that are most relevant to priming, as McRae and colleagues themselves, as well as others, have persuasively argued (McRae & Boisvert, 1998; McRae, de Sa, & Seidenberg, 1997; Lucas, 2000; Hutchison, 2003). Thus a better metric of possible language knowledge is the distributional structure of language itself. In Corpus Study 1, we examine whether language statistics exist that could be used to predict the behavioral results of Ferretti et al.'s Experiment 1.

2.1. Method

2.1.1. Corpus

For both corpus studies, we used a corpus comprised of the entirety of the English online encyclopedia *Wikipedia*, as of October 16, 2006. Each *Wikipedia* subject article was appended into one large corpus. The corpus was cleaned such that all formatting tags and links were removed, leaving only raw text, including punctuation. The 10/16/06 *Wikipedia* corpus contains 5,266,981 unique words, and 518,339,522 total tokens. Beyond its size, the corpus has a number of desirable qualities, including breadth of topic, a relatively jargon-free style, and text produced by a large number of authors.

2.1.2. Analyses

We extracted data from the corpus on the same 24 past-progressive verb forms (the two-word phrase *was X-ing*), 24 past-perfect verb forms (*had X-ed*), and the 24 location nouns that were used by Ferretti et al. (2007). Our analysis used the same design as Ferretti et al., crossing aspect (past progressive, past perfect) and semantic relatedness of the verb phrase-location pair (related, unrelated). Our dependent measure was co-occurrence: the conditional probability of the location noun, given the verb phrase.

To assess co-occurrence probability, we calculated the probability that each aspectual verb form co-occurred within a sentence with its related and with its unrelated location, calculated as a co-occurrence probability of the target location occurring, given the frequency of the verb phrase. For example, the probability of *highway* occurring in a sentence, given that the sentence contained the semantically related verb phrase *was driving*, was 7.93% (there were nine *highway* + *was driving* co-occurrences, out of 126 total occurrences of *was driving* in the corpus). In contrast, the co-occurrence probability of *was driving* and the semantically unrelated location *temple* was 0.79% (one co-occurrence out of the 126 *was driving* occurrences). We analyzed the effects of aspect and relatedness in an ANOVA with items as the random effect.

2.2. Results and discussion

The means and standard errors of the corpus co-occurrence probabilities, as well as the means and standard errors from Ferretti et al.'s behavioral study, are shown in Figs. 1A and 1B, respectively. The co-occurrence probabilities in the corpus closely mirrored the reaction times found by Ferretti et al. In the corpus data, there was an interaction between verb phrase aspect and semantic relatedness ($F_2(1,23) = 6.70, p = 0.016$).² As is clear from Fig. 1A, the nature of this interaction was that the three conditions for which there were no differences in reaction time in Ferretti et al.'s data (past progressive-unrelated location pairs, past perfect-unrelated-location pairs, and past perfect-related-location pairs) were all equally unlikely to co-occur with each other in the corpus, whereas the past progressive-related-location pairs, which yielded the shortest reaction times in Ferretti et al., were more likely to co-occur in sentences than the other three. These results show that the knowledge that Ferretti et al. attributed specifically to world-knowledge event schema is encoded in the co-occurrence patterns of the English language. It is therefore possible that this co-occurrence knowledge alone could be sufficient for explaining the behavioral results that Ferretti et al. attributed to the use of world knowledge. In Corpus Study 2, we extend these lexically-specific results by investigating the extent to which these effects in English are construction-wide relationships, to see if the past progressive verb-location association is a language general pattern, or if the strong past progressive-location association pattern is limited to the particular items used in Ferretti et al.'s study.

3. Corpus Study 2

The purpose of Corpus Study 2 was to investigate the extent to which the language knowledge that links verb aspect and location is specific to particular verb phrase and location pairs, or is in fact a general property across the language as a whole. Corpus Study 2 also allowed us to look for lexically-specific deviations from the construction-wide pattern. This enabled us to develop stimuli in which the collocation statistics of particular words violated the general aspect-location patterns. This in turn allowed us to explore the extent to which comprehension behavior is driven by knowledge of lexically-specific statistics, knowledge of construction-wide statistics, and knowledge of events in the world.

3.1. Method

3.1.1. Stimuli and analyses

In Corpus Study 2, we again used the verbs from Ferretti et al. (2007), but did not restrict ourselves to their set of 24 locations. We calculated each verb's co-occurrence probability with the 10,000 most frequent words in the Wikipedia corpus (calculating separately for each of the verb's two aspectual forms). Next, for each of these 48 verb forms, we compiled a list of the 100 most frequently co-occurring nouns for each verb (which comprised nearly all co-occurrences that occurred more than once; this emphasis on higher co-occurrences is expected to yield more reliable and typical patterns of verb-location usage). We then calculated three statistics for all 48 verbs: the number of locations for each verb in past progressive and past perfect forms, the frequency of the highest co-occurring location in each verb form, and the summed co-occurrence frequency across the top 100 co-occurring nouns. These three measures were designed to assess whether the past progressive-location relationship is a construction-wide pattern, as well as the extent to which there may be deviations from this pattern, both in terms of idiosyncratic and specific verb-location pairings, as well as verbs that deviate from this pattern as a whole.

3.2. Results and discussion

The first analysis addressed whether past progressive and past perfect verbs differ in the number of locations with which they tend to co-occur. Out of the top 100 most frequently co-occurring nouns for

² For this analysis, and all other significance tests comparing conditional probability values, the values were first transformed using the *arcsin* square root transformation, as is appropriate for non-normally distributed proportion data (McDonald, 2014).

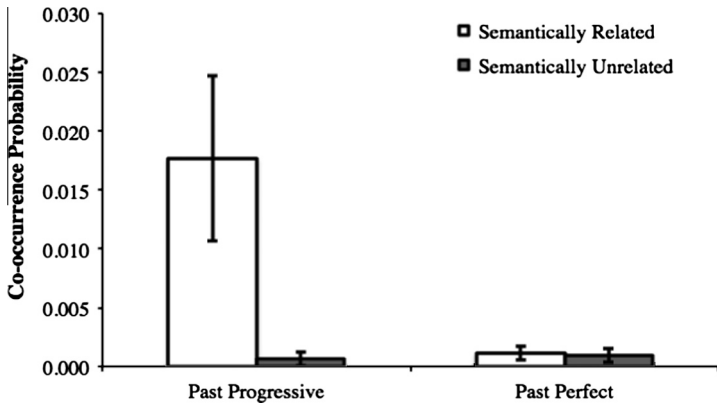


Fig. 1A. The mean co-occurrence probabilities (and standard errors) of the verb phrase-location pairs from Ferretti et al. (2007), calculated within a sentence as the probability of the location given the verb phrase in a 540-million-word Wikipedia corpus.

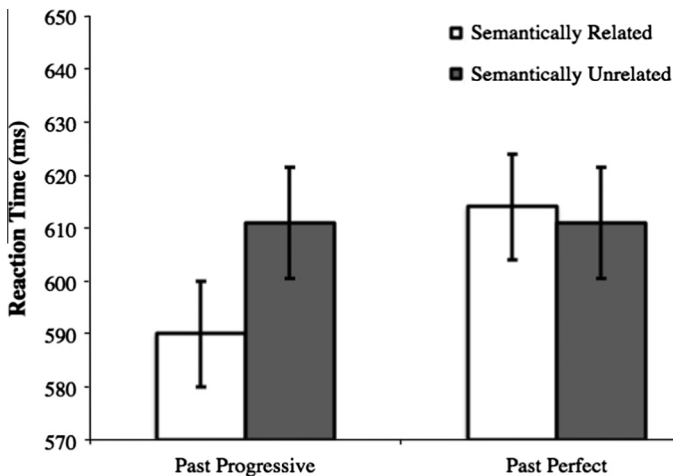


Fig. 1B. The mean reaction times (and standard errors) in the semantic priming naming experiment conducted by Ferretti et al. (2007) (Experiment 1).

each verb, past progressive verbs on average co-occurred with 9.71 locations ($SE = 0.87$), whereas past perfect verbs on average co-occurred with 7.58 locations ($SE = 0.99$), $t(23) = 2.12$, $p < 0.05$. Fig. 2A contains a histogram for past progressive and past perfect verb forms, showing how many of the verbs co-occurred with a small or a large number of different locations. The figure shows that while there is a general tendency for past progressive verb forms to co-occur with a wider variety of locations, six verbs had the opposite pattern (the past perfect form had more location collocates than the past progressive form), and in general the two distributions have considerable overlap.

Our second analysis investigated whether past progressive and past perfect verbs differ in the co-occurrence probability of their most frequent location collocate. There was no significant difference between past progressive verbs ($M = 0.00292$, $SE = 0.00082$) and past perfect verbs ($M = 0.00420$, $SE = 0.00091$), $t(23) = 1.03$, $p = 0.312$.

Our final analysis examined summed co-occurrence frequency for all the location nouns in set of 100 nouns for each verb. Past progressive verbs were more likely to co-occur with locations overall ($M = 0.01146$, $SE = 0.0023$) than past perfect verbs ($M = 0.00719$, $SE = 0.0012$), $t(23) = 2.28$, $p = 0.032$.

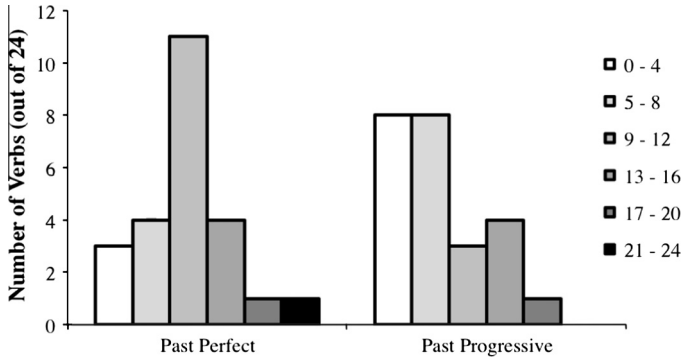


Fig. 2A. Histograms for past progressive vs. past perfect verbs in the Wikipedia corpus, showing the number of verbs that co-occur with different numbers of location words. For example (describing the two leftmost bars), three of the past progressive verbs co-occurred with between 0 and 4 locations, and 4 of the past progressive verbs co-occurred with between 5 and 8 locations.

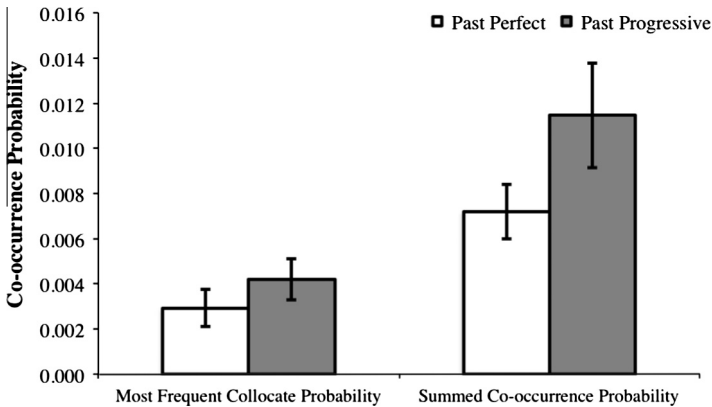


Fig. 2B. The mean co-occurrence probabilities (and standard errors) of the most frequent location collocate, and the mean overall probability of co-occurrence with a location, for both past progressive and past progressive verbs.

The average co-occurrence probability of the most frequent location collocate, as well as the overall probability of co-occurring with location, are shown in Fig. 2B.

In order to be sure these results were not contingent on the Wikipedia corpus, we replicated these corpus analyses using the Corpus of Contemporary American English (COCA, Davies, 2008), a 450-million word corpus from spoken and written genres. The 24 verbs were lower in frequency in the COCA corpus than in the Wikipedia corpus, making the results somewhat noisier, but the general pattern of results was highly similar. Past progressive verbs tended to co-occur with more locations, had higher co-occurrence probability for first collocates (this difference was not reliable in Wikipedia), and had higher summed co-occurrence probability. But as in Wikipedia, there was considerable variance, with many highly co-occurring past perfect-location pairs.

In sum, the Corpus Study 1 showed that the semantically related past progressive verb-location pairs used in Ferretti et al. (2007) co-occur more frequently than their semantically related past perfect verb-location pairs. Corpus Study 2 showed that this result is not restricted to the specific locations used in Ferretti et al.'s study, but rather is a more general pattern in the English language. In addition to the event schema hypothesis suggested by Ferretti et al., these co-occurrence results open the door to the possibility that participants' facilitation to semantically related locations from past progressive verbs (but not past perfect verbs) may arise from prior experience with these words'

co-occurrence patterns. We now turn to behavioral experiments to investigate how these different sources of knowledge may be brought to bear in a variety of comprehension tasks.

4. Behavioral experiments overview

In Experiment 1 (A and B) and Experiment 2 (A and B), we directly tested whether language knowledge or world knowledge (such as event schemas or some other kind of non-linguistic semantic knowledge) better described various semantic effects in participants' language processing and use. To distinguish the various possibilities we have identified in world and language knowledge use, we manipulated four factors. The first three factors were manipulated within each of the four experiments: (1) the semantic relatedness of a particular verb and location noun, (2) the aspectual form of the verb (past progressive vs. past perfect), and (3) the word co-occurrence probability of the verb and location noun, measured in the Wikipedia corpus. The fourth factor, manipulated across experiments, was the extent to which the demands of the task were expected to draw more heavily on language knowledge or on world knowledge.

Experiments 1A and 1B used two tasks that were predicted to force participants to be relatively more reliant on world knowledge than language knowledge. In Experiment 1A, participants made unspeeded, explicit plausibility judgments about verb-locations pairs, and in Experiment 1B participants performed a speeded semantic priming task where they were instructed to decide whether or not the target word was a location. Thus, both tasks involved explicit use of semantic knowledge, and therefore were expected to bias the participants toward the use of world knowledge rather than knowledge about word co-occurrences. In contrast, Experiments 2A and 2B consisted of two tasks in which participants were predicted to be relatively more reliant on language knowledge. In Experiment 2A participants performed a semantic priming task where their only requirement was to name the target word aloud as quickly as possible, and in Experiment 2B participants performed an unspeeded cloze task, completing a sentence with the first word that came to mind. Thus both tasks in Experiment 2 involved the participants performing language-centered tasks, where language knowledge like word co-occurrence information might be expected to play a larger role than world knowledge.

4.1. Stimuli used in all experiments

In the following experiments, we chose stimuli that would help us disentangle language knowledge, specifically aspect-marked verb-location co-occurrences, from knowledge about events in the world. First, we selected 30 high frequency target nouns that are easily identifiable as locations (e.g. *stadium*, *kitchen*, *graveyard*, etc.). For each target location, we then picked three verbs that were used to instantiate a 2×3 design. The first factor was aspect, whether the verb phrase was past progressive (e.g. *was X-ing*) or past perfect (e.g. *had X-ed*). The second factor combined the semantic relatedness and word co-occurrence variables into three levels: (1) semantically related words with high conditional probability of the location given the verb phrase in the 530-million-word Wikipedia corpus, (2) semantically related words with low conditional probability of the location given the verb phrase in the Wikipedia corpus, and (3) semantically unrelated words (which also had low co-occurrence probability). Critically, the verbs that we used were specially chosen because, for a particular location, one semantically-related verb co-occurred frequently with the location in its past progressive form but not its past perfect form; the other semantically related verb co-occurred frequently with the location in its past perfect form but not its past progressive form; the third, semantically unrelated verb did not co-occur frequently with the location in either of its forms. Thus, for each verb form we could investigate independently the behavioral effects of semantic relatedness derivative from world knowledge, word co-occurrence probability, and verb aspect, as well as the interaction of these factors.

For example (shown in Table 1), for the location *supermarket* the verbs *buy* and *pay* are both semantically related in terms of world knowledge. They are both actions that frequently take place in that location, and which are likely to be included as critical aspects of a situation model or event schema for that location. Despite the fact that both *buy* and *pay* are semantically related to *supermarket*, in the

corpus the location *supermarket* was of high probability of occurring given the occurrence of the phrase “*was paying*”, while *supermarket* was of low probability given the occurrence of “*was buying*”. Thus, for the past progressive form we have two phrases that are matched in terms of aspect and in being semantically related to the target, but different in terms of word co-occurrence probability. Likewise, in the Wikipedia corpus *supermarket* was of high co-occurrence probability given the phrase “*had bought*”, but was of low probability given the phrase “*had paid*”. In addition, the words were selected so that the past progressive phrases (e.g. *was paying*) and the past perfect phrases (*had bought*) that were high co-occurrence probability were not significantly different from each other in terms of their probability of co-occurrence, nor were the low co-occurrence probability pairs (e.g. *was buying* and *had paid*). Finally, semantically plausible but unrelated verbs from both aspects were selected as controls (e.g. *was cleaning* and *had cleaned*); the controls were matched so that their co-occurrence probabilities were equal to the co-occurrence probabilities of the semantically related words from the low co-occurrence probability condition. The six verbs for all 30 locations are shown in [Appendix A.1](#).

As in the previous corpus study, to ensure that the items selected were not contingent on idiosyncrasies of the Wikipedia corpus, we validated using the COCA corpus, checking to make sure that the values were significantly different in both corpora. The conditional probability values were in each case higher in the “High Probability” conditions than in the corresponding “Low Probability” conditions, and this did not vary significantly by corpus. The mean co-occurrence values in each condition from each corpus are shown in [Table 2](#). For reference, a co-occurrence score of 0.0041 (the value for the Past Progressive High Co-occurrence condition in the Wikipedia corpus) corresponds to saying that for every 10,000 occurrences of the verb phrase in the corpus, the related locations chosen for the experiments (and shown in [Appendix A.1](#)) co-occurred with that verb phrase on average 41 times. This contrasts with an average of 2 and 1 times in the low probability and unrelated conditions, respectively. Unrelated (or low co-occurrence) items will still co-occur occasionally due to chance, but at much lower rates than in our high probability conditions.

4.2. Predictions

In this section, we identify several possible outcomes of these studies and their implications for accounts of language processing, language use, and semantic representations.

The first hypothetical outcome is that participants will demonstrate only a main effect of semantic relatedness (e.g., in a semantic priming task, showing facilitation for semantically related verb-location pairs, regardless of their co-occurrence probability). In its simplest form, a semantic relatedness effect would not interact with verb aspect or any language co-occurrence statistics. A simple semantic relatedness effect like this may be more likely than in previous studies, given the care that was taken to balance our stimuli for other factors.

A second hypothetical outcome is a replication of the verb aspect \times relatedness interaction that [Ferretti et al. \(2007\)](#) obtained, such that participants show facilitated responses to (or give higher ratings to, etc.) semantically related locations, but only for past progressive items. As we have noted, [Ferretti et al.](#) attributed such data patterns to the online use of “world knowledge event schemas” as a function of specific syntactic frames (see also [Ferretti et al., 2001](#); [McRae, Hare, & Tanenhaus, 2005](#)), but this result is also consistent with a construction-wide effect in which aspect markings modulate the likelihood of locations, independent of particular verb-location co-occurrences. If participants showed facilitation for past progressive items regardless of co-occurrence statistics, this would be evidence for the position outlined by [Ferretti et al.](#); the past progressive verb form activates ongoing event schemas, which then activate their related locations.

A third possibility is that participants will show effects of using word co-occurrence information and no effect of world knowledge (either of a basic semantic-relatedness variety or the more complex pattern predicted by [Ferretti et al.](#)’s event schema hypothesis). This situation would yield main effects of lexical co-occurrence without interactions with other factors. This would be evidence for a language knowledge effect, and would also be evidence that no construction-wide abstract knowledge (about the types of arguments that are likely to co-occur with “*was-X-ing*” or “*had-X-ed*” frames) was being used.

Table 1Design for Experiment 1: The six possible primes for the target location *supermarket*.

Co-occurrence probability	Verb phrase aspect	
	Past progressive	Past perfect
High	<i>was paying</i>	<i>had bought</i>
Low	<i>was buying</i>	<i>had paid</i>
Unrelated	<i>was cleaning</i>	<i>had cleaned</i>

Table 2

Average conditional probabilities for the six conditions, in Wikipedia and in the COCA Corpus.

Co-occurrence probability condition	Verb phrase aspect			
	Wikipedia		COCA	
	Past progressive	Past perfect	Past progressive	Past perfect
High	0.0041	0.0036	0.0012	0.0018
Low	0.0002	0.0001	0.0006	0.0004
Unrelated	0.0001	0.0001	0.0001	0.0002

A fourth alternative is for a more complex use of linguistic information in which both co-occurrence information and construction-wide information affect responses. Such frequency-by-regularity interactions are commonplace in behavioral experiments and computational models of language learning and use (Juliano & Tanenhaus, 1993; Plaut, McClelland, Seidenberg, & Patterson, 1996; Plunkett & Marchman, 1993), including in interpretation of verb-noun relations (Pearlmutter & MacDonald, 1995), and we may well expect to see them here as well if language knowledge is a key driver of behavioral results in certain tasks. If so, then all locations for past progressive verb forms should be facilitated, even if they do not frequently co-occur, because the language-wide tendency is for past progressive-location co-occurrences. This result is also predicted in Ferretti et al.'s account (owing to world knowledge), but the two accounts make different predictions for the pattern of results for the past perfect verb forms. In our account, high frequency past perfect-location pairs should be facilitated, owing to their high co-occurrence in the language, whereas in Ferretti et al.'s account, there is no support in the event schema for facilitation of past perfect-location pairs.³

A final caveat for these predictions concerns potential effects of task. One possibility is that task will have little or no effect, suggesting that language and/or world knowledge is routinely activated in a consistent way, no matter what the participant's task. Alternatively, the participants' task across the four experiments may modulate the degree to which participants rely on language vs. world knowledge, suggesting that language users are able to draw more heavily on one or the other as needed.

5. Experiment 1

5.1. Experiment 1A

In Experiment 1A, we used a task where world knowledge was expected to be very relevant and influential: unspedded, explicit plausibility judgments about relationships in the world. Explicit semantic plausibility judgments arguably comprise the strongest test of whether word co-occurrence

³ In a footnote, Ferretti et al. allow that there may be completed event schemas (activated by had-X-ed frames) that *are* likely to include associations to locations, and thus lead to facilitation (pg. 184 of Ferretti et al., 2007). Thus Ferretti et al. do entertain the possibility that there could be specific exceptions to their general aspect-driven world knowledge perspective. This would be an example of a frequency-by-regularity interaction within event schemas, not within language knowledge, because they suggest that there could be exceptional *events*, not exceptional language statistics or other linguistic information.

probability will influence behavior above and beyond semantic relatedness. In addition, the presence or absence of an interaction of semantic relatedness with aspect will be informative as to the type of world knowledge people might employ when making judgments about the world, and whether or not more complex representations of world knowledge (such as Ferretti et al.'s event knowledge hypothesis) are employed even in explicit, unspeeded tasks like plausibility judgments.

5.1.1. Method

5.1.1.1. Participants. Seventy-eight University of Wisconsin–Madison undergraduates participated in Experiment 1A and received course credit for their participation. All participants were native-English speakers with normal or corrected to normal vision.

5.1.1.2. Materials. The 180 verb phrase–location pairs (30 locations \times 6 experimental conditions) were assigned to six lists, such that each target location and each aspect-marked verb form occurred only once on each list. Each list contained an equal number of past progressive and past perfect verb phrases. In addition, 10 unrelated and low co-occurrence probability filler pairs (5 of each aspect) were added to each list so that the number of related and unrelated pairs (20) was equal on each list.

5.1.1.3. Procedure. Participants sat at a computer and read a list of verb phrase and location pairs. They were instructed, “For each pair, you will see a phrase describing an event or an action, and a noun describing a location. On a 1 to 7 scale, rate how likely is it that the event or action described typically takes place in the location listed. When making this rating, 7 means ‘very likely to take place in this location,’ and 1 means ‘not likely at all’ to take place in this location.” Participants were instructed to type a number 1 through 7 reflecting their opinion on the line next to each pair. Participants were instructed to work at whatever pace they felt comfortable, and that we were not at all interested in how long they took to make their decisions.

5.1.2. Results and discussion

For the plausibility ratings in Experiment 1A, there was a significant main effect of the three-level relatedness condition [$F_1(2,77) = 384.15, p < 0.001$; $F_2(2,29) = 275.85, p < 0.001$], no effect of aspect (all F 's < 1), and no aspect-by-relatedness interaction [$F_1(2,77) = 1.18, p = 0.31$; $F_2(2,29) = 1.02, p = 0.37$]. Means and standard errors for Experiment 1A are shown in Fig. 3. As is clear from the figure, semantically related pairings were rated as highly plausible, regardless of aspect or co-occurrence probability (high and low probability semantically related pairs were not significantly different, all F 's < 1).

In Experiment 1A, using a task that should highly prioritize world knowledge and semantic relatedness information, and using carefully controlled stimulus items that remove the confounds between verb aspect, semantic relatedness, and word co-occurrence probability, we found a simple effect of semantic relatedness and no effect of aspect or word co-occurrence probability. Thus, in a very explicit, unspeeded semantic reasoning task, we found no evidence of sensitivity to word co-occurrence probability, nor of more complex semantic knowledge such as Ferretti et al.'s event knowledge schema activation theory. Beyond the demonstration that this task shows no evidence of effects of language statistics, these results are informative in that they provide normative semantic relatedness ratings for these stimuli: according to this large sample of participants, all of the related verb phrase/location pairs in our study (dark and light gray bars in Fig. 3) were equally “semantically related”. Any differences between high and low co-occurrence probability related pairs that we may find in subsequent studies can therefore be attributed to co-occurrence statistics and not simple effects of semantic relatedness.

5.2. Experiment 1B

In Experiment 1B, we used a second task in which world knowledge was predicted to be relevant and influential. We used a speeded semantic priming task, in which participants made a semantic judgment about a target word (“Is it a Location?”). Like Experiment 1A, this task was designed to privilege world knowledge. However, unlike the explicit rating task, the speeded semantic judgments

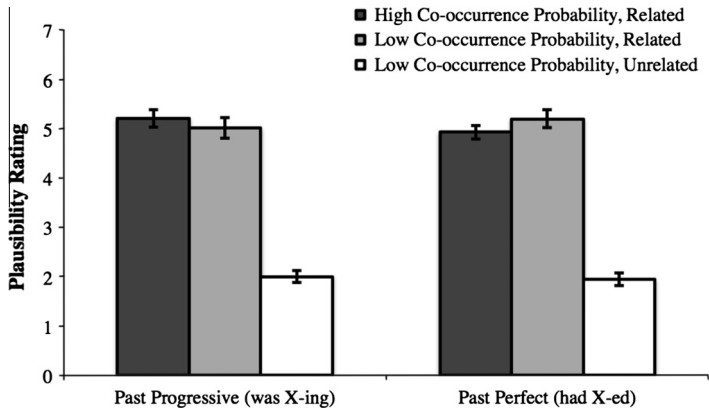


Fig. 3. Participants' mean ratings (and standard errors) of the plausibility of verb phrase – location pairs (e.g. *was buying – supermarket*), as a function of the verb phrase's aspect, the pair's semantic relatedness, and the pair's probability of co-occurring in sentences in the 530-million-word Wikipedia corpus.

allow greater sensitivity for detecting more nuanced properties of semantic relatedness, such as the effects of word co-occurrence probability or the activation of event schemas, if those effects exist. Thus, if the explicit, unspeeded nature of the task in Experiment 1A masked subtler effects, Experiment 1B could be expected to reveal more complex results. However, if task speed was not an issue, then we might expect to find simple semantic relatedness effect, replicating Experiment 1A.

The priming task employed the same timing and display parameters as the one reported by Ferretti et al. (2007), but with two important differences: (1) our items removed the confound between semantic relatedness and word co-occurrence probability, and (2) this experiment made use of a location judgment task rather than the target naming task used by Ferretti et al. This task difference is critical, as the use of semantic judgment tasks in semantic priming tasks may be expected to emphasize the use of semantic knowledge relative to other strategies that participants may employ to solve a task (McRae et al., 1997).

5.2.1. Method

5.2.1.1. Participants. Seventy-eight University of Wisconsin–Madison undergraduates participated in Experiment 1B and received course credit for their participation. All participants were native English speakers with normal or corrected to normal vision. None of the participants had participated in Experiment 1A.

5.2.1.2. Materials. The critical items for Experiment 1B consisted of the 180 verb phrase–location pairs (30 locations \times 6 experimental conditions) from Experiment 1A. These items were assigned to the same six, between-participant lists containing 30 items each, such that each target location and each aspect-marked verb form occurred only once on each list. Each list contained an equal number of past progressive and past perfect verb phrases.

The task for Experiment 1B was a semantic judgment task, where for half of the trials the correct response was “yes” and for half the trials the correct response was “No.” There were 160 total trials for each participant, 80 “yes” trials and 80 “no” trials. The 80 “yes” trials consisted of that participant's list of 30 critical items and 10 semantically unrelated, low-co-occurrence probability filler items that were used to balance semantic relatedness. An additional 40 filler items were added to each list. These filler items were also all semantically unrelated, low-co-occurrence probability verb–location pairs. These additional filler items were added to reduce the relatedness proportion (i.e. the proportion of trials where the prime and target were semantically related) to 0.25, in line with levels typically deemed to be necessary to eliminate strategic processing effects from semantic priming experiments (McNamara, 2005). The “no” experimental stimuli consisted of an additional 80 verb–noun pairs where

the target word was *not* a location. For these filler items, the target word was always a concrete object noun (e.g. *dog* or *hammer*), and for these trials 25% were semantically related, mirroring the “yes” trials so that semantic relatedness could not serve as a cue to the correct response).

5.2.1.3. Procedure. The procedure for Experiment 1B was designed to mimic previous experiments where semantic judgments were used to elicit semantic priming effects (i.e. McRae et al., 1997). Each trial began with a fixation cross in the center of the screen for 250 ms, followed by a prime verb phrase (e.g. “*was paying*”). The prime remained on screen for 200 ms, at which point it was replaced for 50 ms by a pattern mask (&&&&&&&&). Thus, this 250 ms stimulus onset asynchrony helped ensure that there were no strategic effects in the experiment (McNamara, 2004). After the pattern mask disappeared, the target location word appeared on the screen and remained until the participant’s response. Participants were instructed to read the prime and target silently, and then to make a judgment about the target word: whether or not it was a location. In the instructions, participants were given the example *house* for which the correct answer was yes, and participants were given the example *dog* for which the correct answer was No. While a *dog* could in principle be a location (e.g., for *fleas*), our instructions and examples emphasized a conventional sense of “location” as being a location for humans. Participants were told to respond as quickly as possible and indicate “yes” or “no” by pressing the 1 or 2 keys on the keyboard.

5.2.2. Results and discussion

Average accuracy across all trials (i.e. correctly identifying the target as either a location or not a location) was above 90%, and did not vary significantly by condition (all F 's < 1). This eliminates concerns that participants misunderstood the task or conceived of “no” trials (such as *dog*) as possible locations, as well as eliminated any speed-accuracy tradeoff concerns.

We then analyzed reaction times for semantic priming effects. For each participant, only the 30 critical experimental trials were analyzed, corresponding to the 30 critical target items. All trials with reaction times less than 300 ms and greater than or equal to three standard deviations of the mean reaction time were excluded from the analysis (28 out of 2340 trials). Due to the number of between-subjects comparisons and because not all participants saw all of the exact same verb phrase-location pairs, all naming reaction times were z-scored according to each subject’s mean reaction time⁴ (Gelman & Hill, 2006). In Experiment 1B, as in Experiment 1A, there was no main effect of aspect, nor an interaction between aspect and relatedness condition (all F 's < 1). There was a main effect of relatedness significant by participants [$F_1(2, 77) = 3.67, p = 0.03$] and marginally significant by items [$F_2(2, 29) = 3.04, p = 0.063$]. Follow-up tests on the effect of relatedness (collapsing across the verbs’ aspectual form) found that responses to both semantically related high co-occurrence probability targets [$F_1(1, 79) = 4.22, p = 0.043$] and to semantically related low co-occurrence probability targets [$F_1(1, 79) = 4.59, p = 0.035$] were significantly shorter than responses to unrelated targets, and that responses to high and low co-occurrence probability targets were not different from each other. The reaction times and standard errors for Experiment 1B are shown in Fig. 4.

Thus, in Experiment 1A and 1B we found very similar patterns of results: simple effects of semantic relatedness without effects of word co-occurrence probability, or interactions with a verb’s aspect. These results suggest that at least for these stimuli and semantic judgment tasks, the nature of the judgment promotes the use of world knowledge, and time pressure does not modulate this effect.

Further, the failure to replicate the aspect-by-relatedness interaction found by Ferretti et al. using this different task and stimuli may be instructive about the process by which priming did occur in Ferretti et al.’s naming task. As we showed in Corpus Studies 1–2, Ferretti et al.’s items confounded semantic effects due to world knowledge and effects due to word co-occurrence, and as such the aspect-by-relatedness interaction they found in their naming task may have been due to world knowledge, knowledge of co-occurrence in language, or some combination of the two. We pursue this question in Experiment 2.

⁴ The effects reported were not qualitatively affected by z-scoring or by the omission of the outlier trials, though this did affect the p -values of the ANOVAs by small amounts (all reported significant effects never rose above $p = 0.07$ as a result of either of these manipulations).

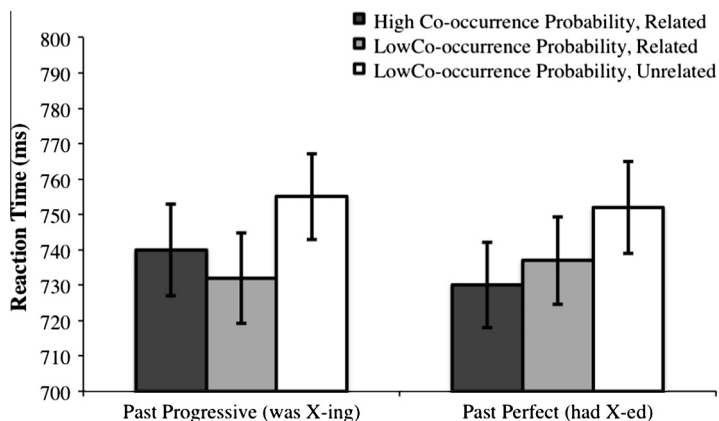


Fig. 4. Participants' mean reaction times (and standard errors) in a semantic priming experiment with a semantic judgment task – as a function of the verb phrase's aspect, the pair's semantic relatedness, and the pair's probability of co-occurring in sentences in the 530-million-word Wikipedia corpus.

6. Experiment 2

Experiments 2A and 2B used the same materials as in Experiments 1A–B but explored tasks that were designed to emphasize the relevance of language knowledge rather than world knowledge. Contrary to a theory that places all of the effects in world event knowledge (i.e. facilitation of semantically related past progressive verb–location pairs, but not semantically related past perfect–location pairs), in these more language-focused tasks in Experiment 2, we predicted that results would be driven by differences in word co-occurrence statistics. Experiment 2A employed the same naming task used by Ferretti et al. (but with our more carefully controlled items). Experiment 2B used a sentence completion task, in order to test if any effects observed in Experiment 2A were due to the speeded nature of the task rather than the language-centric-bias of the task.

6.1. Experiment 2A

6.1.1. Method

6.1.1.1. Participants. Seventy-eight University of Wisconsin–Madison undergraduates participated in Experiment 2A and received course credit for their participation. All participants were native-English speakers with normal or corrected to normal vision. None of the participants had participated in Experiments 1A or 1B.

6.1.1.2. Materials and procedure. The materials and procedure for Experiment 2A were identical to those in Experiment 1B, with the exception of the task instructions. Participants were instructed to read the prime word silently and to read the target word aloud, as quickly and as accurately as possible.

6.1.2. Results and discussion

A research assistant who was blind to the prime condition examined the waveform for each spoken response, discarded all trials in which participants had named the target incorrectly (<1% of the data), and measured speech onset from the waveform, using WaveSurfer software. All trials with reaction times less than 300 ms, and greater than or equal to three standard deviations of the mean reaction time were excluded from the analysis (<1% of the data). The raw reaction times and standard errors for Experiment 2A are shown in Fig. 5. Due to the number of between-participants comparisons, all

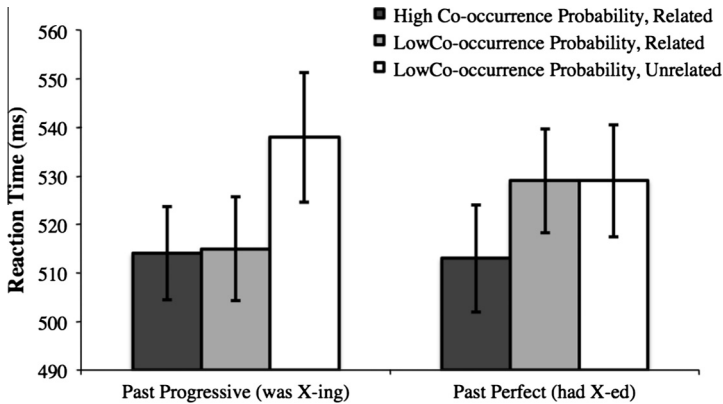


Fig. 5. Participants' mean reaction times (and standard errors) in a semantic priming experiment with a naming task – as a function of the verb phrase's aspect, the pair's semantic relatedness, and the pair's probability of co-occurring in sentences in the 530-million-word Wikipedia corpus.

naming times were z-scored according to each participant's mean reaction time. There was not a significant main effect of aspect, when analyzed by participants or by items (both F 's < 1). There was a significant main effect of relatedness (high co-occurrence vs. low-co-occurrence vs. unrelated) both by participants [$F_1(2,77) = 9.25, p < 0.001$] and by items [$F_2(2,29) = 8.71, p = 0.001$]. There was also an interaction between aspect and relatedness that was significant by participants [$F_1(2,77) = 5.00, p = 0.009$] and by items [$F_2(2,29) = 5.00, p = 0.013$].

Follow-up comparisons showed that for past perfect verbs, participants were faster in the high co-occurrence probability related condition than in the low co-occurrence probability related condition [$F_1(1,77) = 5.96, p = 0.017, F_2(1,29) = 4.28, p = 0.048$] and faster in the high co-occurrence probability related condition than in the unrelated condition [$F_1(1,77) = 4.04, p = 0.048, F_2(1,29) = 4.00, p = 0.055$]. For past perfect verbs, reaction times in the related, low-probability condition and the unrelated condition were not significantly different (F_1 and $F_2 < 1$). In contrast, comparisons for the past progressive condition showed that naming was faster for semantically related verbs regardless of co-occurrence probability: high probability [$F_1(1,77) = 20.88, p < 0.001, F_2(1,29) = 13.07, p < 0.001$] and low probability [$F_1(1,77) = 18.67, p < 0.001, F_2(1,29) = 15.71, p < 0.001$] related verbs were both faster than unrelated verbs, and did not differ from each other (F_1 and $F_2 < 1$).

This pattern of results suggests the use of both item-specific word co-occurrence knowledge and construction-wide knowledge about aspect-location probabilities. Short reaction times in the high co-occurrence conditions (regardless of aspect, shown in the two dark bars) reflect facilitation from past experience with individual verb-location co-occurrences. Short reaction times in the past progressive semantically related but low co-occurrence condition are explainable as a consequence of a frequency-by-regularity interaction, where reaction times reflect the fact that overall, past progressive verb phrases are more likely to co-occur with locations (as shown in Section 3). On this view, both the past progressive *was X-ing* construction and individual verb-location co-occurrences speeded naming of location targets.

The data are not consistent with a simple version of a world knowledge event schema theory, where past progressive (*was X-ing*) frames activate ongoing event schemas that contain locations, whereas past perfect (*had X-ed*) frames activate completed events schemas that do not contain locations. Such a theory does not predict the short reaction times for high co-occurrence probability past perfect verb-location pairs (in addition to not predicting why we found no effect for low co-occurrence probability past progressive pairs in Experiment 1A or 1B). But as Ferretti et al. (2007) note in Footnote 1, it is possible for event schema theories to be more complex; some ongoing event schemas may include locations, whereas for some actions where the completion of the event is very important (such as accomplishments like climb), locations may be a part of the completed event schema as well.

However, even with the addition of exceptional events with atypical schemas, an account that places *all* of the semantic effects within world knowledge schemas cannot easily explain the divergence of results in Experiment 1A, 1B, and 2A. If locations *are* a part of the event schemas for our frequently co-occurring past perfect pairs, then according to an (exclusively) event schema theory they should have differentially activated these locations in Experiment 1, providing higher ratings (in 1A) or faster reaction times (in 1B). And if they are *not* a part of the completed event schemas, then they should not have led to activation and facilitation in the priming experiment in this experiment.

The combined results from Experiments 1A and 1B (where we found semantic relatedness effects that did not interact with aspect in a world knowledge focused task) and Experiment 2A (which showed an effect of word co-occurrence knowledge in a language focused task) argue in favor of a more complex explanation, where both event schemas (or some other sort of world knowledge) and knowledge of language statistics (or some other form of knowledge about language) are both present and being used to varying degrees in response to task demands.

6.2. Experiment 2B

Experiment 2B employed a fragment completion task, in which participants were given an incomplete sentence containing one of the target verb phrases and asked to complete the sentence with the first word or phrase that came to mind. Completion tasks like this one have increasingly been used in comprehension research to assess comprehenders' knowledge of sentence-specific probabilities, that is, how the syntax and meaning of current input (the sentence or discourse fragment) shapes perceivers' expectations for upcoming input (Gennari & MacDonald, 2008, 2009; McRae et al., 1998; see Smith & Levy, 2013, for discussion of completion and corpus data in capturing language knowledge). In contrast to single word completion (Cloze) tasks, the issue is not (or not only) the specific word provided in a completion but rather information at a coarser grain, such as the syntactic structure of the completion or thematic roles conveyed. Completions coded at thematic and syntactic grains have been shown to be reliable correlates of language comprehension difficulty, as measured by reading times (Gennari & MacDonald, 2008, 2009). Given this tradition, the completion task is a good choice for an untimed task that draws on language knowledge (relative to the untimed, semantic plausibility rating task in Experiment 1A), allowing us to investigate whether the effects observed in Experiment 2A were dependent upon the use of a speeded task.

6.2.1. Method

6.2.1.1. *Participants.* Ninety-three University of Wisconsin–Madison undergraduates completed the study for course credit. All identified English as their native language. None of the participants had participated in any of the previous experiments.

6.2.1.2. *Materials.* Each participant saw a list of 90 sentence fragments that contained a unique common proper name and a verb phrase (e.g. “*Mary was visiting. . .*”). The lists were compiled in the following manner. The 180 verb phrases (90 past progressive and 90 past perfect) from the previous experiments were randomly assigned to two lists such that each verb (regardless of aspect) occurred only once in each list, so that each list had an equal number of past progressive and past perfect verb phrases, and such that each list contained an equal number of verbs from the three co-occurrence probability conditions from the previous experiments (related high co-occurrence probability, related low co-occurrence probability, and unrelated conditions).

6.2.1.3. *Procedure.* Participants were seated at individual computers, where they were shown a Microsoft Excel workbook with a list of sentence fragments as described above. The pairings of proper names with verbs and the order of the sentence fragments were randomized for each participant. Excel's autocomplete function was disabled. Participants read instructions asking them to type a natural sounding completion for the fragments, using the first word or phrase that came to mind. Each fragment appeared on its own row in the leftmost column. Participants responded by typing a

completion in the cell immediately to the right of each fragment, starting at the top and working down.

6.2.2. Results and discussion

The completion data were analyzed in two ways. First, the completions were scored for whether or not a location for the given verb was included in the completion. For example, “*Andrea had studied in Spain*” was scored as containing a location, but “*Pamela had studied French for 12 years before visiting France*” was scored as not containing a location, because France is the location for the visiting event, not the studying event that was part of the sentence fragment prompt. Each item was independently scored twice; disagreements were rare and were resolved through discussion. Several verbs had multiple completions for which there was no clear consensus on the presence of a location (e.g., *Julie had marched in the parade*). The rate of unscorable completions did not vary as a function of aspect, and these items were eliminated from further analyses. Once this scoring was completed, there was a count, for each verb, of the total number of locations that were used with each verb. These counts were analyzed as a function of the verb’s aspectual form. In this analysis, past progressive verb forms led to the production of 372 total locations, with an average of 7.59 participants ($SE = 1.15$) listing at least one location in their sentence completions for the fragments containing past progressive form verbs. Past perfect form verbs led to the production of 356 total locations, with an average of 7.27 participants ($SE = 1.20$) giving at least one location in their sentence completions for the fragments containing past perfect form verbs. There was not a significant difference between past progressive and past perfect form verbs in the number of locations listed [$F_2(1, 89) = 1.12, p = 0.29$]. This lack of a difference is not particularly surprising, because these stimulus items were chosen to remove the confound of the tendency of locations to co-occur with these verbs as a function of their aspect. This equivalence in the number of locations produced for these carefully controlled items is further evidence against a strong claim that past progressive syntactic frames should necessarily lead to the direct activation of location words, independent of other factors about those words’ semantic properties.

In a second analysis, the frequency with which particular location nouns were generated in the completion of each verb form fragment was predicted in a multiple linear regression. The regression included two factors: (1) the co-occurrence probability of that particular verb phrase/location pair in the Wikipedia corpus; and (2) the aspectual form of the verb. In other words, we investigated whether participants tended to produce particular location nouns that had high co-occurrence frequency in the corpus. The goal of this analysis was to determine whether language statistics such as co-occurrence probability predict behavior in this linguistic-centric task, and whether the aspectual form of the verb

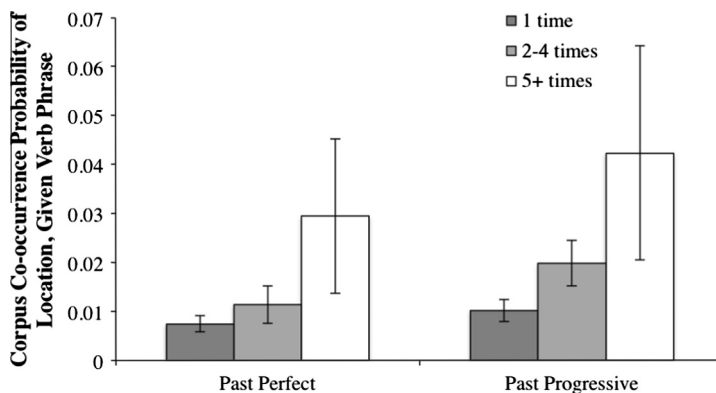


Fig. 6. Corpus co-occurrence probabilities (mean and standard error), as a function of the number of participants who produced that pair in the Experiment 2B.

moderates this predictiveness. In this regression, the verb phrase's co-occurrence probability with the location in the Wikipedia corpus was a significant predictor of the location's production frequency [$t(59) = 5.20, p < 0.001, r = 0.329$] and the aspect of the verb accounted for no additional variance ($p > 0.05$). Fig. 6 shows the average co-occurrence probabilities of verb phrase/location pairs as a function of aspect and how frequently they were produced in fragment completions, showing that as a location became more likely to be produced given a particular verb phrase, its co-occurrence probability in the Wikipedia corpus rose, and that this effect was not significantly moderated by aspect.

The clear co-occurrence effects in Experiment 2B (as well as lack of aspect effects on location production frequency), together with the results of Experiment 2A, support the hypothesis that while people are engaged in linguistic-centric tasks (like sentence completion or reading words aloud), their semantic behaviors are predictable in terms of language statistics, including word co-occurrence probabilities, even to the extent that these co-occurrence probabilities deviate from simple measures of semantic relatedness.

7. General discussion

In this paper, we used two corpus analyses and four behavioral experiments to explore how knowledge of language (about verb-location collocations) and world knowledge (about events and the entities that participate in them) interact during language and semantic tasks. In the corpus analyses, we showed that verbs in past progressive form do indeed co-occur more frequently with locations than do verbs in past perfect form, though with wide variability across verbs. We capitalized on this variation to develop stimuli that disentangled world and language knowledge and used them in four different tasks that varied in the centrality of language knowledge for performing the task. In Experiments 1A and 1B, using two tasks that were expected to bias participants toward using world knowledge and away from using language statistics, we found exactly that: simple effects of semantic relatedness with little or no effect of language statistics. In Experiments 2A and 2B, using two tasks that were expected to bias people toward using language statistics, we found behavior that moved away from the simple effects of semantic relatedness and toward behaviors that mirrored the language statistics.

Our results do not support Ferretti et al.'s (2007) hypothesis that verbs' syntactic frames (i.e. their past progressive or past perfect forms) combine with the verb to differentially activate either an "ongoing event schema" or a "completed event schema" in memory, with ongoing event schemas more likely to lead to the activation of locations in that world knowledge schema, leading to the activation of location words. Across four experiments using our carefully controlled items, we never obtained Ferretti et al.'s pattern of data. Ferretti et al.'s findings thus appear to reflect their use of items whose word co-occurrence probabilities particularly favored the particular interaction that they found, as suggested by Corpus Study 1, in combination with using a task that is likely to draw on language knowledge (a naming task).

Our results argue for strong effects of task demands on the kinds of information that people activate and use. First consider events in the world: when we activate knowledge of events, and think about them in terms of an event that is either ongoing or is completed, there is certainly evidence to suggest that locations (often goal states) are more important in our conceptual representations of ongoing events, compared to completed ones (Athanasopoulos, & Bylund, 2013; Bylund & Jarvis, 2011; Madden & Zwaan, 2003; Magliano & Schleich, 2000). As Glenberg (1997) put it in his seminal paper "What Memory is For" – where he outlined the embodied cognition perspective – objects and events in the world have certain affordances, and these affordances dictate how memories are structured and used. On this view, it could be said that thinking about an ongoing event "affords" thinking about the location in which that event occurred.

Our results suggest it is a mistake, however, to extend these world-event relations to online language use, where a different set of affordances exist, and comprehension processes appear to be oriented to expectation and interpretation as a function of past language experience. Specifically for the aspect-location phenomena we have been considering, there is a contrast between events and

their linguistic representation: although thinking about a particular ongoing event may afford thinking about its location, *talking about* an event may not so strongly afford talking about its location. The communicative and pragmatic constraints involved in talking about an event might well lead to considerably different affordances for a verb than for its referential event. This perspective thus captures “what language is for” (or, at least, one of the things that language is for). Likewise, divergences between language and event representations may also arise for differences in languages’ structures. A number of recent studies have found that the link between event representations and goal locations differ as a function of how a language encodes aspect and related properties of verbs’ constructions (Athanasopoulos & Bylund, 2013; Bylund & Jarvis, 2011). Together, these results and observations suggest that while Ferretti et al. were correct to note the relationship between events and locations, this relationship turns out to be more complex than they hypothesized, with task demands, construction-wide, and individual lexical effects all interacting to determine the resultant behavior.

7.1. Implications for theories of semantic memory

Our results bring into focus the ways in which semantic knowledge combines knowledge about meaning that is derived from world knowledge (knowledge about the typical locations in which various events take place) and from knowledge about meaning that is derived from language experience (such as knowledge of verb–location co-occurrences in language, as well as general tendencies for past progressive verb forms to co-occur with locations). Experiments 1A and 1B provide evidence for the unsurprising conclusion that semantic knowledge is not exclusively derived from knowledge about verbs’ patterns of co-occurrence in language, insensitive to world knowledge about events. Similarly however, Experiments 2A and 2B provide evidence that semantic knowledge is not exclusively derived from world knowledge about events, insensitive to distributional knowledge about verbs. Together these studies indicate that task demands can shift the extent to which people rely on their prior experience with the world or with language.

This proposal need not be seen as at odds with schema theories of semantic knowledge more generally, such as that proposed by McRae and colleagues in a number of papers (Ferretti et al., 2001, 2007; Hare et al., 2009; McRae et al., 1997, 1998). What we suggest is that this notion of schemas be extended to include lexical information as a complimentary part of a schema representation. Rather than proposing that particular morphosyntactic frames (like “was X-ing”) activate specific “ongoing” event knowledge schemas, encountering a particular verb phrase (such as “was cheering”) will have the potential to activate both world knowledge related to *cheering* events as well as lexical knowledge derived from experiences talking about, hearing about, or reading about *cheering* events. These different kinds of associations will be activated and used in a manner that is in accordance with their relevance to the task in which the person is engaged. Or to put another way, those different kinds of information will be activated to the extent that they are afforded by (or useful to) the action in which the person is engaged. Such a schema representation could, but need not, build in an explicit representational distinction between words and their referents, as differences between these two types of input would likely emerge as a function of the vastly different roles these different types of input play. Thus, such a theory would retain the benefits of schema theories over traditional, more structured models of semantics (see Rogers & McClelland, 2004, for a review). At the same time, it would retain the advantages of these earlier structured models (as well as earlier two-stage models of sentence processing) in noting important differences between semantic knowledge derived from language and semantic knowledge derived from the world. But unlike those earlier models, these differences could emerge due to differences in use, rather than as built-in parts of the system.

This proposal is consistent with a series of recent studies discussing the relationship between world knowledge and language knowledge – proposals that have made the point that representational systems combining both of these types of knowledge have distinctive advantages. Riordan and Jones (2010) modeled learning to semantically categorize words, based on either world knowledge (operationalized in terms of a word’s semantic features) or language knowledge (operationally

defined in terms of words' distributional similarity in a large corpus of naturalistic language) and found that knowledge derived from language *or* from the world could be used equally well to learn to categorize common nouns. Riordan and Jones concluded that this was evidence of redundancy in perceptual and language experience. However, they also noted that while models based on semantic features and linguistic distributional similarity perform equally well in this task, each model's performance seems to make use of slightly different information. Riordan and Jones note, for example, "distributional models seem to give more weight to information about actions, functions, and situations, and less to information about direct perception related to objects (e.g., texture, internal properties, etc.)" (p. 37).

Similarly, [Andrews, Vigliocco, and Vinson \(2009\)](#) tested three Bayesian models of semantic priming, one using exclusively word-specific knowledge (again operationally defined in terms of the distributional similarity of words in terms of their usage across different contexts), one using exclusively world knowledge (again operationally defined in terms of the similarity of words' semantic feature overlap), and one model that combined both types of information. Consistent with the current results, the model containing both language knowledge and world knowledge was the best fit to semantic priming data, though we would also expect that the relative importance of language and world knowledge could vary with task demands. Andrews et al. note that this advantage of the composite model is not a function of merely having *more* information. Instead they argue that the composite model performs better because it is combining distinct, complimentary sources of information.

Andrews et al. cast this argument about how language knowledge and world knowledge are combined in a Bayesian learning framework, arguing that an intelligent learning system should be able to use the redundant information across the two types of input to align those representations, and then combine the information that only comes from language or only comes from the world into one composite representation. More recently, [Johns and Jones \(2012\)](#) have proposed a formal learning model that does exactly this. They use a computational model to show that it is a relatively simple exercise in inductive inference to draw correct conclusions about the perceptual properties about a word's referent – even when one has not experienced that word's referent in the physical world – based on the word's distributional similarity to other words about which we do have world knowledge.

7.2. Implications for online language comprehension

Our findings, especially the fact that the usage of language knowledge and world knowledge seems to be modulated by the appropriateness of that knowledge to the task, support the hypothesis that the representation of semantic knowledge and its use during language processing is an exquisitely context-sensitive system. This system can be thought of as an extension of the system outlined in multiple constraint satisfaction approaches to online comprehension (e.g., [MacDonald et al., 1994](#); [Tanenhaus & Trueswell, 1995](#)), with the additional theoretical claim being that the system is not just a mixed bag of constraints that interact in a task or context-insensitive manner. The current research does not explain how comprehenders come to deploy different knowledge sources in different task environments, but a reasonable first hypothesis is that this flexibility is learned over time, such that there is a bias to weigh more heavily the information that has previously proved to be more informative in similar contexts. These hypotheses merit additional investigation.

Similarly, our studies have investigated only a small piece of language comprehension, within rather constrained tasks, and it will be important to extend these findings to other comprehension measures, such as reading times and eye movements in complete sentences or texts. The findings presented here about reliance on language statistics, together with other results in sentence processing that have found that reading times are strongly affected by word co-occurrences and other linguistic patterns ([Levy, 2008](#); [McDonald & Shillcock, 2003](#); [Smith & Levy, 2013](#)), suggest that online comprehension of aspect-location relationships would be better accounted for by word co-occurrences, perhaps in combination construction-wide language statistics, than by world knowledge.

This point has some interesting implications for psycholinguistic methods, regarding attempts to match experimental sentences for semantic information in studies where some other factor (such as ambiguity or syntactic complexity) is being investigated. A common method of stimulus control is to collect normative data on the materials, either on the plausibility of the ideas expressed in the sentence (intended to reflect real world knowledge, e.g. [Ferretti & McRae, 1999](#)) or by using sentence completions or corpus analyses to assess potentially relevant language statistics (e.g., [McRae et al., 1998](#)). Both types of measures correlate with reading times, but given the weight of language statistics in our results for language tasks, we can speculate that completion or corpus data may provide a superior estimate of the knowledge that affects online reading times. Indeed, given the very strong relationship between past linguistic experience and reading times ([Smith & Levy, 2013](#)), a very strong position is that plausibility norms correlate with reading times merely by virtue of the fact that world knowledge (assessed by plausibility norms) is usually correlated with the language knowledge that actually drives much of online processing. And indeed, word co-occurrence and distributional similarity measures can be used to predict plausibility quite well ([Willits, Duran, D'Mello, & Olney, 2007](#)). We consider this a possible outcome in some circumstances, but we do not endorse this speculation as a general characterization of all comprehension processes. Quite the contrary, our broader claim throughout has been that the comprehender's task demands will affect the degree of reliance on language vs. world knowledge. There remains substantial work to be done to better understand this division of labor, but one outcome of the present studies is that we can reject the notion that a semantic effect in sentence comprehension necessarily reflects the application of world knowledge.

The current results also speak to a point noted in the introduction, that early theories of sentence processing hypothesized the late use of world knowledge in comprehension, on the view that the comprehension system could not compute this information rapidly enough to affect online comprehension. Our findings suggest an alternate solution to the need for rapid computation, that online comprehension processes may be driven more by language statistics than by statistics about the world. This emphasis on language statistics for language processing suggests a sort of functional encapsulation of the language comprehension system, but it is not the architectural modularity of [Frazier \(1987\)](#) and similar accounts. Instead the greater reliance on language statistics may emerge from prior experiences in which this information has proved effective in aiding comprehension processes. We suspect that these phenomena are quite general, where rapidly computed statistical contingencies substitute for much more complex relationships with which they are correlated. Some examples include language producers' use of their own knowledge as a proxy for inferences about what their audience knows vs. doesn't know ([Barr & Keysar, 2007](#)), comprehenders' interpretation of prosody as indicating syntactic or semantic properties of the utterance ([Schafer, Speer, & Warren, 2005](#)), and studies in which orthographic statistics signal grammatical category of a word, apparently even before the word is recognized ([Dikker, Rabagliati, Farmer, & Pytkkanen, 2010](#)). As these examples suggest, we expect that language knowledge extends well beyond the word co-occurrences that were the focus of inquiry here. Thus while we have investigated task-based reliance on language vs. world knowledge in one domain, aspect-location information, we expect that the phenomena of relying on an efficient subset of all available knowledge, even within the world knowledge and language knowledge domains, is a very broad phenomenon.

Acknowledgments

This research was funded by NSF Grant BCS-1123788, a grant from NIDCD (F31DC009936), and by NIH Training Grant 5T32HD049899 to the University of Wisconsin–Madison. We would like to thank Rachel Sussman, Ken McRae, Mark Seidenberg, and Jessica Montag, Chuck Clifton, and three anonymous reviewers for useful comments and discussion, as well as many members of the Language and Cognitive Neuroscience Lab for their assistance running participants in the experiments and analyzing data.

Appendix A.1. Stimulus items used in Experiments 1 and 2

Location noun	High probability for progressive form, low probability for past perfect form	High probability for past perfect form, low probability for progressive form	Unrelated (low probability for both forms)
Airport	Arriving/arrived	Scheduling/scheduled	Burying/buried
Apartment	Moving/moved	Arranging/arranged	Drowning/drowned
Ballroom	Singing/sung	Recording/recorded	Recuperating/recuperated
Bank	Advising/advised	Depositing/deposited	Cheering/cheered
Bar	Drinking/drank	Serving/served	Marrying/married
Base	Operating/operated	Deploying/deployed	Educating/educated
Beach	Swimming/swam	Drowning/drowned	Stealing/stolen
Cell	Hanging/hung	Imprisoning/imprisoned	Auditioning/auditioned
Cemetery	Praying/prayed	Burying/buried	Anchoring/anchored
Church	Preaching/preached	Marrying/married	Raiding/raided
Clinic	Recovering/recovered	Visiting/visited	Racing/raced
Club	Performing/performed	Hosting/hosted	Graduating/graduated
College	Teaching/taught	Enrolling/enrolled	Paying/paid
Court	Prosecuting/prosecuted	Testifying/testified	Camping/camped
Forest	Hunting/hunted	Blossoming/blossomed	Prosecuting/prosecuted
Highway	Driving/driven	Turning/turned	Hanging/hung
Mountain	Training/trained	Sliding/slid	Arriving/arrived
Museum	Showing/shown	Displaying/displayed	Drifting/drifted
Nightclub	Dancing/danced	Assaulting/assaulted	Stowing/stowed
Ocean	Drifting/drifted	Sinking/sunk	Assaulting/assaulted
Park	Pitching/pitched	Camping/camped	Experimenting/experimented
Port	Shipping/shipped	Docking/docked	Hosting/hosted
Prison	Awaiting/awaited	Escaping/escaped	Docking/docked
River	Fishing/fished	Anchoring/anchored	Drinking/drank
School	Learning/learned	Educating/educated	Losing/lost
Street	Racing/raced	Running/ran	Enrolling/enrolled
Studio	Rehearsing/rehearsed	Mixing/mixed	Journeying/journeyed
Supermarket	Buying/bought	Paying/paid	Cleaning/cleaned
Theatre	Auditioning/auditioned	Staging/staged	Advising/advised
Vault	Stealing/stolen	Storing/stored	Perishing/perished

Note, in all experiments, the stimulus verbs were presented with the grammatically appropriate auxiliary (*was* for past progressive verbs, *had* for past perfect verbs).

References

- Andrews, M., Vigliocco, G., & Vinson, D. P. (2009). Integrating experiential and distributional data to learn semantic representations. *Psychological Review*, *116*, 463–498.
- Amato, M. S., & MacDonald, M. C. (2010). Sentence processing in an artificial language: Learning and using combinatorial constraints. *Cognition*, *116*(1), 143–148.
- Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language*, *62*, 67–82.
- Athanasopoulos, P., & Bylund, E. (2013). Does grammatical aspect affect motion event cognition? A cross-linguistic comparison of English and Swedish speakers. *Cognitive Science*, *37*, 286–309.
- Barr, D. J., & Keysar, B. (2007). Perspective taking and the coordination of meaning in language use. In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed., pp. 901–938). New York: Academic Press.
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*, 577–660.
- Bloomfield, L. (1933). *Language*. New York: Henry Holt.

- Bylund, E., & Jarvis, S. (2011). L2 effects on L1 event conceptualization. *Bilingualism: Language and Cognition*, 14, 47–59.
- Collins, A. M., & Loftus, E. F. (1975). A spreading activation theory of semantic processing. *Psychological Review*, 82, 407–428.
- Dikker, S., Rabagliati, H., Farmer, T. A., & Pykkänen, L. (2010). Early occipital sensitivity to syntactic category is based on form typicality. *Psychological Science*, 21, 629–634.
- Ferretti, T. R., Kutas, M., & McRae, K. (2007). Verb aspect and the activation of event knowledge. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 33, 182–196.
- Ferretti, T. R., & McRae, K. (1999). Modeling the role of plausibility and verb-bias in the direct object sentence complement ambiguity. In *Proceedings of the Twenty-First Annual Conference of the Cognitive Science Society* (pp. 161–166). Hillsdale NJ: Erlbaum.
- Ferretti, T. R., McRae, K., & Hatherell, A. (2001). Integrating verbs, situation schemas, and thematic role concepts. *Journal of Memory and Language*, 44, 516–547.
- Firth, J. R. (1957). A synopsis of linguistic theory, 1930–1955. In *Philological society (Great Britain) (Ed.), Studies in linguistic analysis* (pp. 1–32). Oxford, England: Blackwell.
- Fisher, C., Gleitman, L. R., & Gleitman, H. (1991). On the semantic content of subcategorization frames. *Cognitive Psychology*, 23, 331–392.
- Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 559–586).
- Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, 6, 291–325.
- Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.
- Gennari, S. P., & MacDonald, M. C. (2008). Semantic indeterminacy in object relative clauses. *Journal of memory and language*, 58(2), 161–187.
- Gennari, S. P., & MacDonald, M. C. (2009). Linking production and comprehension processes: The case of relative clauses. *Cognition*, 111(1), 1–23.
- Gentner, D. (1981). Verb semantic structures in memory for sentences: Evidence for componential representation. *Cognitive Psychology*, 13, 56–83.
- Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences*, 20, 1–55.
- Grice, P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics* (Vol. 3). New York: Academic Press.
- Hale, J. (2001). A probabilistic Earley parser as a psycholinguistic model. In: *Proceedings of the second meeting of the North American Chapter of the Association for Computational Linguistics on Language technologies* (pp. 1–8). Association for Computational Linguistics.
- Hare, M., Jones, M., Thomson, C., Kelly, S., & McRae, K. (2009). Activating event knowledge. *Cognition*, 111, 151–167.
- Harris, Z. S. (1957). Co-occurrence and transformation in linguistic structure. *Language*, 33, 283–340.
- Hutchison, K. A. (2003). Is semantic priming due to association strength or featural overlap? A micro-analytic review. *Psychonomic Bulletin & Review*, 10, 785–813.
- Johns, B. T., & Jones, M. N. (2012). Perceptual inference from global lexical similarity. *Topics in Cognitive Science*, in press.
- Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.
- Jones, M. N., Kintsch, W., & Mewhort, D. J. K. (2006). High-dimensional semantic space accounts of priming. *Journal of Memory and Language*, 55, 534–552.
- Juliano, C., & Tanenhaus, M. K. (1993). Contingent frequency effects in syntactic ambiguity resolution. In *Proceedings of the fifteenth annual conference for the cognitive science society*.
- Katz, J. J., & Fodor, J. A. (1963). The structure of semantic theory. *Language*, 39, 170–210.
- Kiss, G. R. (1973). Grammatical word classes: A learning process and its simulation. *Psychology of Learning and Motivation*, 7, 1–41.
- Landau, B., & Gleitman, L. R. (1985). *Language and experience: Evidence from the blind child*. Cambridge, MA: Harvard University Press.
- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The Latent Semantic Analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review*, 104, 211–240.
- Lany, J., & Saffran, J. R. (2010). From statistics to meaning: Infants' acquisition of lexical categories. *Psychological Science*, 21, 284–291.
- Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, 106, 1126–1177.
- Louwerse, M. M. (2008). Embodied relations are encoded by language. *Psychonomic Bulletin and Review*, 15, 838–844.
- Lucas, M. (2000). Semantic priming without association: A meta-analytic review. *Psychonomic Bulletin & Review*, 7, 618–630.
- Lund, K., Burgess, C., & Audet, C. (1996). Dissociating semantic and associative word relationships using high-dimensional semantic space. In *Proceedings of the eighteenth annual conference of the cognitive science society* (pp. 603–608). Mahwah, NJ: Erlbaum.
- MacDonald, M. C. (1993). The interaction of lexical and syntactic ambiguity. *Journal of Memory and Language*, 32, 692–715.
- MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101, 676–703.
- McDonald, J. H. (2014). *Handbook of biological statistics* (3rd ed.). Baltimore, Maryland: Sparky House Publishing.
- McNamara, T. P. (2005). *Semantic priming: Perspectives from memory and word recognition*. Psychology Press.
- Madden, C. J., & Zwaan, R. A. (2003). How does verb aspect constrain event representations? *Memory & Cognition*, 31, 663–672.
- Magliano, J. P., & Schleich, M. C. (2000). Verb aspect and situation models. *Discourse Processes*, 29, 83–112.
- Maratsos, M., & Chalkley, M. (1980). The internal language of children's syntax: The ontogenesis and representation of syntactic categories. In K. Nelson (Ed.), *Children's language* (Vol. 2). New York: Gardner Press.
- McDonald, S. A., & Shillcock, R. C. (2003). Eye movements reveal the on-line computation of lexical probabilities during reading. *Psychological Science*, 14, 648–652.
- McRae, K., & Boisvert, S. (1998). Automatic semantic similarity priming. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 24, 558–572.

- McRae, K., de Sa, V. R., & Seidenberg, M. S. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General*, 126, 99–130.
- McRae, K., Hare, M., & Tanenhaus, M. K. (2005). Meaning Through Syntax is insufficient to explain comprehension of sentences with reduced relative clauses: Comment on McKoon and Ratcliff (2003). *Psychological Review*, 112, 1022–1031.
- McRae, K., Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language*, 38, 283–312.
- Mintz, T. H. (2003). Frequent frames as a cue for grammatical categories in child directed speech. *Cognition*, 90, 91–117.
- Monaghan, P., Chater, N., & Christiansen, M. H. (2005). The differential role of phonological and distributional cues in grammatical categorization. *Cognition*, 96, 143–182.
- Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language*, 17, 357–374.
- Nelson, D. L., McEvoy, C. L., & Schreiber, T. (1999). *University of South Florida word association, rhyme and word fragment norms*. <<http://cyber.acomp.usf.edu/FreeAssociation/>>.
- Patson, N. D., & Warren, T. (2014). Comparing the roles of referents and event structures in parsing preferences. *Language, Cognition and Neuroscience*, 29, 408–423.
- Pearlmutter, N. J., & MacDonald, M. C. (1995). Individual differences and probabilistic constraints in syntactic ambiguity resolution. *Journal of Memory and Language*, 34, 521–542.
- Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, 103, 56–115.
- Plunkett, K., & Marchman, V. (1993). From rote learning to system building: Acquiring verb morphology in children and connectionist nets. *Cognition*, 48, 1–49.
- Reali, F., & Christiansen, M. H. (2007). Processing of relative clauses is made easier by frequency of occurrence. *Journal of Memory and Language*, 53, 1–23.
- Redington, M., Chater, N., & Finch, S. (1998). Distributional information: A powerful cue for acquiring syntactic categories. *Cognitive Science*, 22, 425–469.
- Reichenbach, H. (1947). *Elements of symbolic logic*. New York: Free Press.
- Riordan, B., & Jones, M. N. (2010). Redundancy in linguistic and perceptual experience: Comparing distributional and feature-based models of semantic representation. *Topics in Cognitive Science*, 3, 303–345.
- Rogers, T. T., & McClelland, J. L. (2004). *Semantic cognition: A parallel distributed processing approach*. Cambridge, MA: MIT Press.
- Rumelhart, D. E., & Levin, A. (1975). A language comprehension system. In D. A. Norman & D. E. Rumelhart (Eds.), *Explorations in cognition* (pp. 179–208). San Francisco: Freeman.
- Schafer, A. J., Speer, S. R., & Warren, P. (2005). Prosodic influences on the production and comprehension of syntactic ambiguity in a game-based conversation task. In M. Tanenhaus & J. Trueswell (Eds.), *Approaches to studying world situated language use: psycholinguistic, linguistic and computational perspectives on bridging the product and action tradition* (pp. 209–225). Cambridge: MIT Press.
- Smith, N. J., & Levy, R. (2013). The effect of word predictability on reading time is logarithmic. *Cognition*, 128, 302–319.
- Spivey-Knowlton, M., & Sedivy, J. (1995). Parsing attachment ambiguities with multiple constraints. *Cognition*, 55, 227–267.
- Swinney, D. (1979). Lexical access during sentence comprehension: (Re) Consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, 18, 645–659.
- Tanenhaus, M. K., & Trueswell, J. C. (1995). Sentence comprehension. In J. L. Miller & P. D. Eimas (Eds.), *Handbook of perception and cognition. Speech, language and communication* (Vol. 11, pp. 217–262). San Diego, CA: Academic Press.
- van Dijk, T. A., & Kintsch, W. (1983). *Strategies in discourse comprehension*. New York: Academic Press.
- Wasow, T. (1997). End-weight from the speaker's perspective. *Journal of Psycholinguistic research*, 26(3), 347–361.
- Willits, J. A., Duran, N. D., D'Mello, S. K., & Olney, A. (2007). Distributional statistics and thematic role relationships. In *Proceedings of the 19th annual meeting of the cognitive science society*.
- Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, 123, 162–185.