CrossMark

Applied Psycholinguistics **36** (2015), 1193–1221
doi:10.1017/S0142716414000113

Processing collocations in a second language: A case of first language activation?

BRENT WOLTER
Idaho State University

JUNKO YAMASHITA
Nagoya University

Received: April 22, 2013 Accepted for publication: December 24, 2013

ADDRESS FOR CORRESPONDENCE
Brent Wolter, Department of English and Philosophy, Idaho State University, 921 South 8th Avenue, Stop 8056, Pocatello, ID 83209-8056. E-mail: woltbren@isu.edu

ABSTRACT
This study investigated the possible influence of first language (L1) collocational patterns on second language (L2) collocational processing. A lexical decision task was used to assess whether collocational patterns acceptable in the L1 but not the L2 would still be activated when processing language entirely in the L2. The results revealed no such activation. Furthermore, L2 speakers did not produce accelerated processing for control collocations that were acceptable in the L2 but not the L1. Based on these findings, we put forth some theoretical suggestions regarding recent research indicating accelerated processing for congruent over incongruent collocations. Finally, our NS control group revealed some unexpected tendencies that cannot be easily accounted for with our current understanding of L1 language processing.

This paper addresses the processing of collocations in a second language (L2). Collocations are lexical patterns that are dictated more by convention within the language than by grammatical or purely semantic restrictions. In English, for example, it would be natural to talk about someone with *narrow eyes* who *broke a promise*, but any skilled speaker would notice the comparative novelty of being told of a person with *thin eyes* who *ripped a promise.* In Japanese, however, the exact opposite would be true. Collocations are often viewed as a subcategory of formulaic language (see, e.g., Hoey, 2005, Sinclair, 1991; Wray, 2002), and knowledge of collocations, and other formulaic patterns, is felt to be an important component for proficient language use (e.g., Hoey, 2005; Nation, 2001; Schmitt & Carter, 2004; Wray, 2002, 2008).

Despite their apparent pervasiveness in language, collocations are notoriously difficult to define. There is no agreed upon definition for what constitutes a

© Cambridge University Press 2014 0142-7164/14 $15.00

https://doi.org/10.1017/S0142716414000113 Published online by Cambridge University Press

collocation (for discussions, see Barfield & Gyllstad, 2009; Gyllstad, 2007; Nesselhauf, 2003). However, there have been two distinct approaches historically. The first is often referred to as the phraseological approach (see Cowie, 1998; Howarth, 1998). Under this approach, for a string of words to be considered a true collocation, at least one of the words in the string must be used in a semantically nontransparent way. Thus, a phrase like *read a book* would be considered a free combination rather than a collocation, while *read one's mind* would be considered a collocation because it does not involve the physical act of reading in the core sense of the word. The second approach is largely rooted in corpus linguistics, and it involves relying on statistical measures to determine how frequently certain word patterns co-occur in a large corpus or corpora (e.g., Sinclair, 1991). Various measures exist, but the basic idea behind them is that they provide an estimation of the extent to which certain words "attract" one another in authentic language use.

Naturally, there are advantages and disadvantages to either approach. Regardless of the approach one chooses to apply in defining collocations, however, they are, as mentioned above, still considered by many researchers to be an important component of one's overall language competence. Although we do not yet fully understand the range of factors that might affect L2 collocational processing, the research so far has highlighted some revealing tendencies. To begin with, more frequent L2 collocations tend to be processed more quickly than less frequent L2 collocations (Wolter & Gyllstad, 2013; for a discussion of binomials, see also Siyanova-Chanturia, Conklin, & van Heuven, 2011). In addition to frequency, *congruency* in collocation has also been found to affect processing. Congruent collocations are those that, when translated word for word, are felicitous in both the first language (L1) and the L2. An example of a congruent collocation in English and Japanese would be *strong wind(s).* These English words would typically be translated into Japanese as *tsuyoi* and *kaze*, respectively, and the resulting collocation *tsuyoi kaze* is both natural and very similar in meaning to the English expression. An example of an incongruent collocation would be the English expression *high wind(s)*. Although it works in English, when translated word for word into Japanese (*takai kaze*) it results in an expression that would be perceived as either unnatural or at least quite distinct in meaning from its English counterpart. The research to date has indicated that congruent collocations tend to be processed more quickly than incongruent collocations (Wolter & Gyllstad, 2011, 2013; Yamashita & Jiang, 2010), although it should be pointed out that these categories are complicated to some extent because words do not always have simple and straightforward translations (e.g., the Japanese word *semai*, which can reasonably be translated into either *narrow* or *small* in English).

Theoretically speaking, the aforementioned frequency effects in L2 collocational processing are not difficult to account for, particularly among higher proficiency learners. Although it is almost universally accepted that L2 acquisition is different from L1 acquisition in some important and fundamental ways, frequency effects can easily be explained by a variety of theoretical accounts. These include the large number of theoretical positions that fall more or less under the broad rubric of usage-based grammars (e.g., Bybee, 2007; Bybee & Hopper, 2001; Goldberg, 2006; Langacker, 1987, 1988; MacWhinney, 2000; and Tomasello, 2000, 2003). A

https://doi.org/10.1017/S0142716414000113 Published online by Cambridge University Press

major premise of these accounts is the idea that more frequent encounters lead to greater entrenchment of larger chunks of language that in turn leads to more rapid activation in subsequent encounters. It also fits well with age of acquisition/order of acquisition (AoA/OoA) effects (Carroll & White, 1973; see detailed discussion below), because more frequent language patterns are more likely to be encountered early on in one's learning experience than less frequent patterns. Finally, these findings are consistent with spreading activation accounts of the lexicon (Collins & Loftus, 1975) and the idea that efficiency in language processing is attained through probabilistic reasoning derived from previous language experiences (see e.g., McDonald & Shillcock, 2003a, 2003b). Under this view, when a language user encounters a given word in language input, it activates other words that have commonly co-occurred with it in their prior language experiences. This in turn accelerates recognition of residually activated words once they are actually encountered.

Finding a theoretical explanation to account for the apparent advantages afforded by congruency in collocational processing, however, is less straightforward. Although much has been written over the years in respect to the L1's influence on L2 acquisition, the majority of this research has focused on grammar (e.g., Jiang, Novokshanova, Masuda, & Wang, 2011; Marinis, Roberts, Felser, & Clahsen, 2005; Papadopoulos & Clahsen, 2003) or lexis defined as single words (e.g., Jiang, 2004; Sunderman & Kroll, 2006; Wang, Koda, & Perfetti, 2003). Much less attention has been given to the L1's influence of language acquisition of lexis at larger grain sizes. In response to the lack of a theoretical framework to explain the L1's influence in L2 collocational processing, Wolter and Gyllstad (2011, p. 444) present the possibility of dual activation of collocational connections (i.e., simultaneous and parallel activation of congruent collocational patterns in both the L1 and the L2). This was based on the results of a primed lexical decision task (LDT) in English that showed significantly faster reaction times (RTs) to target words that were primed with words that formed congruent collocations (for L1 Swedish speakers) over target words that were primed with words that formed incongruent collocations. The authors speculated that this accelerated RT might be explained by L1 collocational knowledge running in the background to assist the processing of L2 collocations. In other words, they conjectured that it is possible that when one is presented with an L2 word, the word activates not only known L2 collocates but also L1 collocates. In the case of congruent collocations, this would essentially amount to a replication of activation which would feed back to the visual input and allow for more rapid recognition than would be the case for incongruent collocations (which must be processed solely in terms of the L2 network and would not benefit from L1 knowledge). Though there are some important distinctions, this view bears similarity to theoretical accounts of empirical results regarding speed of recognition for (single-word) cognates (see Lemhöfer & Dijkstra, 2004).

The key question that arises here is whether this view can be reconciled with existing perspectives regarding activation across a speaker's languages. Doing so requires validation of two key assumptions: (a) that there is spreading activation from a word to its collocates and (b) that there is a way in which L1 collocational knowledge can affect processing of L2 collocations. In respect to the first

https://doi.org/10.1017/S0142716414000113 Published online by Cambridge University Press

assumption, although there is a wealth of research investigating priming between semantically related words (for an overview, see e.g., McNamara, 2004) collocational priming is an issue that has received less empirical attention.[1] Nonetheless, there is a growing body of evidence indicating that words prime their collocates as well. In applied linguistics research, Wolter and Gyllstad (2011) found evidence of collocational priming in a primed lexical decision task for both L1 and L2 speakers. Similarly, Durrant and Doherty (2010) found evidence of priming for high-frequency collocations among native speakers (NSs) of English. The results of both of these studies are consistent with psycholinguistic perspectives that view language processing, particularly in respect to reading, as conditioned in part by probabilistic reasoning. McDonald and Shillcock (2003a, 2003b), for example, explored the effects of transitional probabilities (i.e., the corpus-based likelihood that word *y* will follow immediately after word *x*) on eye-tracking measures. They found that the most parsimonious account for the data was attained by combining transitional probabilities with other measures that have also been shown to affect processing (e.g., word frequency). From this they concluded that "the brain is able to draw upon statistical information in order to rapidly estimate the lexical probabilities of upcoming words" (2003a, p. 648). In brief, there seems to be sufficient evidence from these different research perspectives to accept the first assumption spelled out above and accept the likelihood that there is definitely spreading activation from a word to its commonly occurring collocates.

The second assumption is that there is a way for L1 collocational knowledge to affect the processing of L2 collocations. The influence of the L1 on L2 language development, particularly in cases of later L2 acquisition, seems so well established that it may seem reasonable to simply accept this assumption at face value. However, doing so would be tantamount to ignoring a wealth of empirical data that illuminates the complexity of the problem, specifically in respect to online language processing. As noted above, Wolter and Gyllstad (2011) speculated that one explanation for accelerated recognition of congruent collocations was due to parallel activation of the same collocational pattern in the L1. In order for this to happen, it would presumably require an L2 word to stimulate its L1 translation equivalent. Upon receiving this stimulation, the L1 word would presumably stimulate the L1 collocate in concert with the L2 word stimulating its L2 collocate. However, this explanation faces some serious challenges from cross-language priming studies. In particular, the majority of the research to date has suggested that cross-language lexical priming is largely asymmetrical. That is, although there appears to be good evidence indicating lexical priming from the L1 to the L2, there is far less evidence of lexical priming from the L2 to the L1 (for extensive summaries, see, e.g., Brysbaert & Duyck, 2010; Schoonbaert, Duyck, Brysbaert, & Hartsuiker, 2009). In view of this, the parallel activation explanation must be viewed with a good deal of suspicion.

Nonetheless, it remains that when the frequency of collocations is held constant, congruent collocations tend to be recognized more rapidly than incongruent collocations. This means we need to find an alternative explanation to account for this phenomenon. One possibility has to do with how L2 lexical knowledge is acquired by L2 learners. Drawing on the work of Levelt (1989), Jiang (2000) has proposed a three-step process for L2 lexical acquisition. This process assumes a

https://doi.org/10.1017/S0142716414000113 Published online by Cambridge University Press

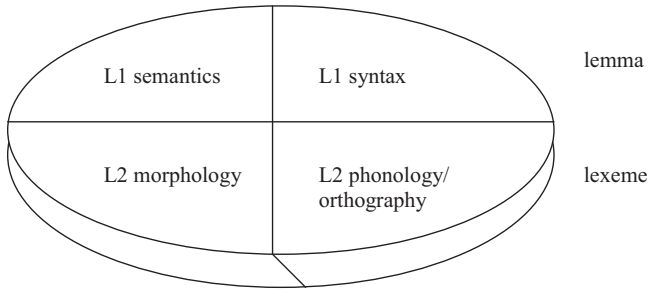


Figure 1. Jiang's model of fossilized second language lexical knowledge. Adapted from *Speaking: From Intention to Articulation*, by W. J. M. Levelt, 1989, Cambridge, MA: MIT Press. Copyright 1989 by MIT Press. Adapted with permission.

representation of lexical items that consists of two levels, the *lemma* level and the *lexeme* level (Figure 1). As can be seen in Figure 1, the lemma level incorporates semantic and syntactic information about the word, while the lexeme level includes morphological and phonological/orthographic information. In Jiang's view, which is based on a comprehensive review of the research literature, the first step in L2 lexical acquisition consists of creating an L2 entry that is linked to a corresponding L1 word. The L1 word in turn provides access to the underlying concept. At this stage, only L2 phonological and/or orthographic knowledge is stored in the L2 entry, and no direct connection exists between the L2 word and the underlying concept. Use of the L2 word is mediated by the corresponding L1 word. Through repeated activation of the L2 word, Jiang contends, the bond between the L2 word and its L1 counterpart is strengthened until the second stage is reached. At this stage, the learner amalgamates information that is copied from the L1 at the lemma level with L2 information at the lexeme level to create an entry that is appropriately L2-like in terms of its morphology and phonology/orthography but which remains L1-like in respect to semantics and syntax. It is essentially a hybrid entry consisting of both L1-based and L2-based knowledge. The third stage is achieved through further exposure to the L2 and involves gradually replacing the L1-based knowledge at the lemma level with more appropriate L2-based knowledge to create an entry that is "very similar to a lexical entry in L1 in terms of both representation and processing" (Jiang, 2000, p. 53).

However, it appears that the last stage is the exception rather than the norm. Jiang presents extensive anecdotal and empirical evidence that suggest that for the majority of learners, fossilization occurs at the second stage "even when extensive contextualized input is available" (2000, p. 54). If we accept Jiang's arguments, then it appears that it is common for L2 learners to have L2 lexical entries with L2-based knowledge at the lexeme level but with lemma knowledge that has been copied directly from the L1. If this is the case, and particularly if L2 lexical entries contain information copied from the L1 at the lemma level, then one could envision a scenario in which collocational knowledge also gets copied over from the L1 to the L2 entries. If one were to accept Firth's famous proclamation that "you shall know a word by the company it keeps!" (1957,

https://doi.org/10.1017/S0142716414000113 Published online by Cambridge University Press

p. 179), then this view becomes axiomatic. Even if one were to adopt a more moderate perspective on the interplay between semantics and collocation, it is not hard to envision a situation in which L1 collocational knowledge could find its way into an L2 lexical entry through this process of L2 lexical acquisition. Furthermore, there is evidence to support the view that the L1 has a strong and lasting influence on L2 collocational choices. Although we are quick to point out that some research (in particular Kellerman, 1979) has shown that L2 learners are reluctant to make conscious judgments about the transferability of collocations (particularly figurative collocations) into the L2, other studies have shown that L2 learners do demonstrate a tendency to transfer assumptions from L1 collocational patterns to the L2 in productive language use (e.g., Bahns, 1993; Bahns & Eldaw, 1993; Fan, 2009; Gabrys-Biskup, 1992; Granger, 1998; Jiang, 2009; Nesselhauf, 2003). Furthermore, we need to highlight the important difference between the sorts of conscious, selective processes that are involved in language production and the psychological, and ultimately subconscious, processes that are involved in language processing. In other words, a Japanese learner of English as an L2 might have learned that she or he should avoid the collocation *rip a promise* when using English productively. However, this does not mean that the lexical entry for the English word *rip* does not still include a collocational link between *rip* and *promise* that has been copied from the L1. It also does not mean that an NS of Japanese who has learned English as an L2 does not experience activation of *promise* upon encountering the word *rip* through the process of spreading activation described above. It has simply become a matter of suppressing what is intuitively and subconsciously activated versus what is publicly produced.

Other explanations exist. Wolter and Gyllstad (2013) put forth the possibility that the processing advantage observed in congruent collocations might be attributable to acquisitional advantages that ultimately manifest themselves in processing behaviors. More specifically, they invoke two closely related theories known as the AoA and OoA effects (Carroll & White, 1973). The basic idea behind these theories is that words (and other forms of knowledge) that are acquired earlier are learned "better" than words that are acquired later. This in turn results in more rapid activation of the earlier acquired words over later acquired words. In carefully controlled empirical studies, Izura and her colleagues (Izura & Ellis, 2002, 2004; Izura et al., 2011) found that OoA could account for differences in RTs for recently learned L2 words on a picture-naming task. Furthermore, they concluded that these effects were more important than other lexical factors that are also believed to influence speed of recognition, including "cumulative frequency, frequency trajectory and imageability" (Izura et al., 2011, p. 32). Izura et al. (2011) evaluated two competing theoretical accounts for the AoA/OoA effect, the s*emantic hypothesis* (Brysbaert, Wijnendaele, & De Deyne., 2000; Ghyselinck, Lewis, & Brysbaert, 2004; Van Loon-Vervoom, Van Ham, & Van der Koppen, 1998) and the *mapping hypothesis* (Ellis & Lambon Ralph, 2000; Lambon Ralph & Ehsan, 2006; Monaghan & Ellis, 2002), and concluded that the most probable explanation came from the mapping hypothesis. The mapping hypothesis is rooted in connectionist models of learning, and it assumes that earlier learned items have a greater influence on the final structure of a network than later learned items, specifically if the earlier learned items continue to be encountered in input and

https://doi.org/10.1017/S0142716414000113 Published online by Cambridge University Press

especially once the network loses plasticity and reaches a relatively stabilized state.

The question that arises here is whether the mapping hypothesis (and AoA/OoA) can also be applied to L2 collocational acquisition as well. More specifically, is there any reason to assume that congruent collocations should be acquired earlier than incongruent collocations? The answer would seem to lie in assuming that congruency in collocation can result in positive L1 transfer that leads to earlier acquisition. Positive transfer occurs when an L2 learner makes valid inferences about linguistic structures in the L2 through reference to similar structures in their L1, and it has been shown to have considerable explanatory power in respect to L2 acquisition for a variety of linguistic features (see, e.g., Gass, 1988; Kasper, 1992). There is no reason to assume that it might not also apply to collocations. Under this view, the L1 collocational pattern would not necessarily be copied directly in the L2 network structure (as was proposed in the explanation based on Jiang's description presented above), but rather the existence of a corresponding L1 pattern would provide the learner with some ready-made assumptions regarding L2 collocations. Upon encountering confirmatory input of congruent collocations in the L2, these assumptions would tend to result in more rapid activation of these congruent collocations over incongruent collocations, the latter of which would not have a corresponding assumption and as such would presumably require more exposures before acquisition occurred. This difference in acquisition order would in turn result in more rapid activation of congruent collocations over incongruent collocations over the long term as predicted by the mapping hypothesis, particularly in cases where initial exposures to a congruent and incongruent collocation occurred at roughly the same time in the learner's input. In short, congruency would tend to encourage earlier acquisition of congruent collocations that would in turn result in a greater influence of congruent collocations over incongruent collocations of the same frequency in the final network structure. Nonetheless, it would still be possible for higher frequency incongruent collocations to achieve earlier acquisition than less frequent congruent collocations if the exposures to the incongruent collocations were sufficient to lead to their acquisition before the same level of sufficiency were reached for the congruent collocations. In simpler terms, it all boils down to a matter of when something is acquired. Although congruency would predispose congruent collocations to acquisition through fewer exposures, this does not preclude the possibility that incongruent collocations could be acquired earlier if the input were sufficient.

This view is consistent with existing research findings into L2 collocational processing that have shown an advantage for congruent collocations over incongruent collocations within, but not necessarily across, frequency bands. Wolter and Gyllstad (2013) found that congruent collocations were uniformly judged as acceptable more quickly than incongruent collocations within each of the eight collocational frequency bands they tested. However, when comparisons were made across the frequency bands, higher frequency incongruent collocations tended to be judged as acceptable more quickly than lower frequency congruent collocations. This finding was similar to that of Yamashita and Jiang (2010), who also observed a processing advantage for congruent over incongruent collocations. Yamashita and Jiang (2010) concluded from their findings that congruency led to a difference

https://doi.org/10.1017/S0142716414000113 Published online by Cambridge University Press

not only in processing speed but also in respect to the influence of frequency of input vis-à-vis congruent versus incongruent collocations:

> We can imagine that, because congruent collocations can be accepted on the basis of their L1 counterparts, the frequency of encounters may not be as important for congruent collocations as it would be for incongruent ones in order to be understood and stored in the L2 lexicon. (p. 663)

Thus far we have two theoretical explanations for the observed accelerated processing of congruent over incongruent collocations. Both explanations seem capable of accommodating the observed differences in the empirical data. However, to the best of our knowledge, there are no studies to date that can provide insights into which account seems more plausible. Therefore, the main aim of this study is to test one of the key assumptions of the first description (the lemma-based explanation) to obtain empirical evidence that will either confirm or refute it as a plausible explanation. Note that the current study does not directly evaluate congruent collocations. Rather, the current study seeks to determine whether collocations that are legitimate in the L1 but *not* the L2 are still activated when processing word combinations in the L2 (e.g., *kill voice*, which works in Japanese but not in English). Evaluating such word combinations should allow us to determine if there is activation of plausible L1 collocations that have no congruent form in the L2, even while processing language entirely in the L2. If activation is found, then it would tend to suggest that collocational knowledge is copied into the lemma level of the lexical entry and in turn to argue in favor a lemma-based explanation for accelerated recognition of congruent collocations. If no such activation is found, however, it would suggest that copying of collocational knowledge into the L2 lemma is unlikely and instead lead us to conclude that another account, such as AoA/OoA effects, may provide a more viable explanation.

In addition, the current study seeks to extend previous research into this area in two ways. The first way is by incorporating L2 learners of various proficiencies to see if proficiency can be linked to any differences in processing tendencies. Although previous research has indicated that congruency affects processing even at higher levels of proficiency (Wolter & Gyllstad, 2011, 2013; Yamashita & Jiang, 2010), we have only limited empirical evidence for the possible decrease of this influence with gains in proficiency (Yamashita & Jiang, 2010). We might expect, as many researchers do, that there is more L1 influence at lower levels of proficiency than at higher levels. For this reason, we decided to include two levels of learners in the present study. The second extension of previous research is in respect to the word class of the collocations and whether this also has a bearing on processing. In some accounts of language systems, such as the aforementioned usage-based grammars, descriptive systems such as word class are thought to have limited explanatory power in accounting for how language systems develop. Instead, the key is frequency of input and regularity of the patterns that achieve some sort of symbolic significance. Tendencies and patterns that can be described through reference to more traditional grammars, and the sort of word class distinctions these entail, are subservient to patterning. In the words of Langacker, "Grammar consists of patterns for creating symbolic assemblies" (2000, p. 21). Therefore, in an effort to probe into the issue of whether different types of collocations lead

https://doi.org/10.1017/S0142716414000113 Published online by Cambridge University Press

to differences in processing, the current study also incorporated adjective–noun (AN) and verb–(object) noun (VN) collocations.

METHOD

Item development

Because the goal of this study was to gauge the activation of the L1 collocational patterns when processing collocations in an L2, we needed to work with groups of nonnative speakers (NNSs) who shared a common L1 and develop items around their L1. For this study we used L1 Japanese speakers (see below), so all items were developed with L1 Japanese speakers in mind. The items fell into three critical conditions: (a) English translations of collocations that were acceptable in Japanese but not English (J-only items, e.g., *far eye, buy anger*), (b) collocations that were acceptable in English but not Japanese (E-only items, e.g., *low speed, catch breath*), and (c) noncollocational items used for baseline RTs to gauge the relative RTs for both the J-only or E-only items.

The initial list of potential items was developed by the second author, who is an NS of Japanese with near nativelike proficiency in English. To ensure that all words would be known to the participants, only high frequency words were used.[2] Existence of all J-only items and nonexistence of the word for word translation of E-only items were verified using a Japanese language corpus.[3] After confirming that the items were either acceptable in Japanese (J-only) or not acceptable in Japanese (E-only), items from both types were submitted to further scrutiny using the Corpus of Contemporary American English (COCA; Davies, 2008). The J-only items were checked to verify that they were not common in this (English-based) corpus, and the E-only items were checked to make sure they were found. For the J-only items, a limited number of occurrences in the corpus was allowed provided that there was not a large, positive mutual information (MI) score associated with the item. MI scores are used to calculate the tendency for words to co-occur with other words in the corpus, and corpus linguists have often used a minimum MI score threshold of 3 as indicative of significant co-occurrence (see, e.g., Hunston, 2002). Given the tremendous size of the COCA, some J-only items did register a small number of occurrences. However, by only accepting items that had an inverse MI score (which indicates that the co-occurrence of one word in proximity to the other word is less likely than in other contexts), we were able to confirm that these were likely due to idiosyncratic rather than significant co-occurrences. In contrast, the goal for the E-items was just the opposite. We wanted to confirm that these did represent corpus-verifiable items, and therefore we only included items that had an MI score of at least 3.0. All MI scores were obtained using nonlemmatized word forms. The mean MI score for all E-only items was 5.38, with a low score of 3.30 (for the item *lucky guess*) and a high score of 7.78 (for the item *free speech*).[4]

The noncollocational baseline items consisted of random recombinations of the words used in the J-only and E-only collocations. This was done to ensure that the lexical frequency of the individual words would be the same between the baseline items and the collocational items. All recombined items were checked against a corpus in the same manner used for the J-only items. If any occurrence did

https://doi.org/10.1017/S0142716414000113 Published online by Cambridge University Press

occur in the corpus, the MI scores were checked to make sure they were inverse. In cases where the recombinations produced positive MI scores, the process was repeated. Through several repetitions, we eventually were able to obtain a list of 105 baseline items that had either no attestation in the corpus or an inverse MI score. Nonetheless, we could not fully eradicate positive MI scores, and in the end we decided to retain three items with positive but low MI scores. These items included *busy teeth* (MI = 0.14), *severe energy* (0.57), and *volunteer phone* (0.28). In total, then, the test consisted of 54 E-only items, 54 J-only items, and 108 baseline items for a total of 216 real-word items.

In addition to these real word items, we also constructed a list of 216 items that included nonce words. Because the study used a LDT that presented both words on the screen simultaneously (see below), we opted for an approach that placed nonce words in the first word (i.e., adjective or verb) position half the time and in the second word (i.e., noun) position the other half of the time. All nonce words were taken from the ARC nonword database (Rastle, Harrington, & Coltheart, 2002) and were matched with the real words in terms of length. As with the baseline items, each nonce word item was paired with a word that had been used in a collocational item in either the J-only or E-only condition. This meant that each real word appeared on the test three times, though in different conditions. To illustrate, for the E-only item *lucky guess*, the adjective also appeared in the baseline item *lucky learner* and the nonce word item *lucky rhurnts*. Similarly, the noun *guess* appeared in the baseline item *dark guess* and the nonce word item *roardge guess*. Repeating the words in the various conditions rather than using different words allowed us to offset any differences in word frequency and word length, both of which are characteristics that are known to affect processing speed in RT measures. By repeating the words, we could be assured that these did not vary across the conditions. In contrast, it did introduce another potential confounder in that it likely lowered the activation threshold for words that had been seen in a different condition (or conditions) earlier in the task. To address this, all items were presented to each participant in an individually randomized order. Thus, a participant was as likely to see a word first in one condition as she or he was to see it in the other two conditions, and it is almost certain that any advantages gained from seeing a word a second or third time evened out both within the individual participant's test and across all of the participants as a whole.

Finally, there is the issue of AN and VN items. An equal number of AN and VN collocational items were developed for both the J-only and the E-only categories. To make sure the items in the two E-only groups were comparable, MI scores were obtained using information from COCA. For AN items, MI scores were calculated on the assumption that the noun would follow directly after the adjective, so MI scores were calculated using a right span of one word after the adjective. For VN items, however, MI scores were calculated under the assumption that for most items the noun might be preceded by an intermediary word like a determiner. Therefore, these were calculated using a right span of up to two words after the verb. Furthermore, in order to ensure comparability across the two types, particularly for the E-only items, the mean corpus frequency for the items in both word classes was kept comparable. The mean, log-adjusted frequency for all the AN E-only items was 5.44 while the VN E-only items was 5.33.[5] An independent

https://doi.org/10.1017/S0142716414000113 Published online by Cambridge University Press

Table 1. *Summary of test items means (with log adjusted frequencies)*

	J-Only		E-Only		Baseline	
Item Type	AN	VN	AN	VN	AN	VN
Length	10.9	10.5	10.4	10.3	10.7	10.4
Adj./verb freq.	10.4	10.6	10.7	10.6	10.5	10.6
Noun frequency	10.9	10.7	10.6	10.6	10.7	10.6
Collocational freq.	—	—	5.4	5.3	—	—

Note: Length refers to the total number of letters for both words. AN, adjective–noun; VN, verb–noun.

samples *t* test revealed no significant difference in the (log-adjusted) frequencies of the items in the two-word classes, $t\,(52) = 0.39$, $p = .698$. A complete summary of the items used in the test is provided in Table 1.

Participants

The participants consisted of one group of English NSs and two groups of English NNSs. The NS group consisted of 26 undergraduates and 1 graduate student at an American university. The NNSs were all NSs of Japanese divided into an advanced group (HNNS group for higher NNS group) and an intermediate group (LNNS group for lower NNS group). The HNNS group was composed of 25 graduate students while the LNNS group consisted of 25 undergraduate students. All NNS participants were studying at the same university in Japan. All participants answered a questionnaire that included questions about their age, vision, and dexterity. The NNSs were also asked to provide information about their English learning background and their (self-reported) proficiency in English. Table 2 summarizes the participants' background information. The NNS groups did not differ in terms of the mean age when they started studying English (both around age 11), but the HNNS group had studied English for a longer period of time (7 vs. 12 years): $t\,(29.39) = -6.3$, $p < .001$. Seventeen students in the HNNS group had experiences of living in English-speaking countries for longer than 1 month, whereas none in the LNNS group had such an experience. The average vocabulary size estimated by Eurocentres Vocabulary Size Test (Meara & Jones, 1990), administered either immediately before or after the experiment, was also significantly different between groups, with the HNNS group having larger vocabulary size (5,558 vs. 4,384), $t\,(48) = -3.46$, $p = .001$.

Administration

The test consisted of a double LDT. An LDT was chosen because we felt it was one of the most effective means of tapping into processing of the J-only collocations. More specifically, our underlying assumption was that if collocational links are

https://doi.org/10.1017/S0142716414000113 Published online by Cambridge University Press

Table 2. *Biographical data for participants (standard deviations)*

Group	*N*	Age	Dexterity (R/L/Both)	Sex (M/F)	Self-Report Proficiency Scores[a]				LOR[b] (months)	EVST
					Speak.	Listen.	Read.	Writ.		
NS	25	26.6 (9.1)	23/2/0	8/17					—	—
HNNS	25	25.2 (3.1)	20/3/2	14/11	4.8 (2.5)	6.2 (2.2)	6.5 (1.7)	5.4 (2.3)	17.3 (28.3)	5557.8 (1330.0)
LNNS	25	19.1 (0.70)	21/4/0	17/8	3.1 (1.1)	3.7 (0.84)	4.7 (1.2)	4.2 (1.5)	0	4383.6 (1056.3)

Note: LOR, Length of residence; EVST, Eurocentres Vocabulary Size Test; NS, native speaker; HNNS, higher nonnative speaker; LNNS, lower nonnative speaker.
[a]1 = *none*, 10 = *nativelike*.
[b]LOR in English speaking countries. Periods shorter than 1 month are disregarded (i.e., counted as none).

copied from L1 lexical entries into L2 lexical entries at the lemma level, then we should expect to still see accelerated activation of these L1-based collocations in contexts where judgments regarding appropriateness in the L2 are not required. This in turn could provide a measure of support for the lemma-based description of L2 collocational development described above. If participants had been asked to judge the qualities of the items in an acceptability judgment task, for example, then J-only items may have had an inhibitory effect rather than facilitating effect. In other words, although it may have allowed us to observe the influence of L1-based collocational knowledge, it would not have allowed us to determine whether this knowledge could account for the accelerated processing observed in previous research on congruent collocations.

There are many ways in which a LDT involving two words can be administered. Jones and Estes (2012) identified three main types: the continuous LDT, the standard LDT, and the double LDT. In the continuous LDT, the two words are presented one at a time and respondents are asked to respond to each word separately. In the standard LDT, a prime (the first word) is presented and then removed prior to the presentation of a target (the second word), but test takers are only asked to respond to the target. In the double LDT, the prime and target are presented simultaneously, and test takers are asked to respond to both words with a single response. Jiang (2012) also identifies three types of LDT, although with some variation in his categorical labels. He suggested that LDTs consist of (a) visual, (b) auditory, and (c) double.

In the current study, we ultimately decided to use a double LDT following the procedure used in Meyer and Schvaneveldt's (1971) initial seminal study on semantic priming, as well as numerous other studies on lexical processing (Chwilla & Kolk, 2002; Fischler, 1977; Kirsner, Brown, Abrol, Chadha, & Sharma, 1984; Manelis & Tharp, 1977; Myerson, Hale, Chen, & Lawrence, 1997; Perea & Rosa, 2002) or even on collocational processing specifically (Ellis, Frey, & Jalkanen,

https://doi.org/10.1017/S0142716414000113 Published online by Cambridge University Press

2009). The most common arrangement of the letter strings in the double LDT is to present both strings in the middle of the screen simultaneously with the first word appearing directly above the second word (Jiang, 2012; see also McDonough & Trofimovich, 2009), so this is the procedure we used in this study as well. The main reason for opting for the double LDT approach over competing variations of the LDT was that it allowed us to skirt the sometimes contentious issue of determining what would be an appropriate stimulus onset asynchrony (SOA) for test takers at different levels of proficiency. SOA refers to the amount of time that is allowed to pass between the initial presentation of the first word and the subsequent presentation of the second word in either a continuous or standard LDT. It is a methodological issue that has attracted much empirical research (as well as debate) because shorter SOAs are thought to lead to automatic (i.e., subconscious) processing, while longer SOAs are thought to lead to strategic (i.e., conscious) processing (for overviews, see, e.g., Lucas, 2000; McNamara, 2004). For our purposes, we were not overly concerned with whether the test takers were tapping into automatic or strategic processes. We were, however, centrally concerned that all the test takers had adequate time to process both words before a response was given. It was mainly for this reason that we ultimately decided on the double LDT format.

The LDT was administered using DMDX software developed at Monash University and at the University of Arizona by K. I. Forster and J. C. Forster. After completing the biographical questionnaire, the test began with a practice session of 20 items followed by the 432 test items (216 real word items, 216 nonce word items). A break was provided after the practice session and in the middle of the test. As noted above, all items were presented in an individualized randomized order as determined by DMDX in order to offset any repetition effects that may favor a particular condition over another. The testing procedure was as follows: First, the test takers were presented with a blank screen for 250 ms. After this, the two-letter strings were presented. These remained on the screen until either a response was made or the item timed out at 3000 ms. At this point, the process started over with the next item.

Participants were not instructed to attend to any collocational properties in the items, nor was the possibility of a connection between the words mentioned at any time. Rather, the participants were instructed to simply press the "yes" key (right control) if both words were real English words and the "no" key (left control) if either word was not a real English word. RTs and error rates (ERs) were logged by DMDX. For the NS group and the HNNS group, the testing process usually took less than 20 min. For the LNNS group, the process on average took just over 20 min.

Results

Two NS participants' data were lost because of a technical error. Therefore, the final analyses were based on data from the remaining 25 NS participants, the 25 HNNS participants, and the 25 LNNS participants. As is typical for LDT analyses, the data were analyzed both in terms of RTs and ERs. Before analyzing the RT data, steps were taken to ensure the integrity of the data by reducing the

https://doi.org/10.1017/S0142716414000113 Published online by Cambridge University Press

Table 3. *Mean log RTs (standard deviations) and ERs for AN and VN items combined*

	RT (*SD*)			Priming Effect		ER (%)		
Group	J-Only	E-Only	Base	J-Only	E-Only	J-Only	E-Only	Base
NS	798 (249)	779 (246)	811 (259)	13	32	1.0	0.9	1.2
HNNS	882 (251)	855 (248)	874 (249)	−8	19	2.1	1.7	1.4
LNNS	992 (350)	968 (351)	986 (346)	−6	18	1.5	2.0	2.1

Note: RTs, Response times; ER, error rate; AN, adjective–noun; VN, verb–noun; NS, native speaker; HNNS, higher nonnative speaker; LNNS, lower nonnative speaker.

Table 4. *Mean log RTs (standard deviations) and ERs for AN items*

	RT (*SD*)			Priming Effect		ER (%)		
Group	J-Only	E-Only	Base	J-Only	E-Only	J-Only	E-Only	Base
NS	803 (258)	771 (237)	803 (255)	0	32	0.6	1.0	1.5
HNNS	895 (261)	858 (247)	875 (247)	−20	17	3.4	1.9	1.9
LNNS	1014 (359)	970 (335)	999 (349)	−15	29	1.0	1.5	2.1

Note: RTs, Response times; ER, error rate; AN, adjective–noun; NS, native speaker; HNNS, higher nonnative speaker; LNNS, lower nonnative speaker.

influence of outliers. To begin with, we eliminated all erroneous responses as well as responses that were accurate but took less than 350 ms. Erroneous responses included all "no" responses to real word items, as well as timeouts. Next we log adjusted all remaining RTs and calculated cutoff points using a threshold of ±2 *SD* computed for each participant individually for all real word RTs and trimmed any responses that fell above or below this threshold. This resulted in the loss of 4.1% of the LNNSs' data, 4.6% of the HNNSs' data, and 5.1% of the NSs' data. Table 3 shows the mean results across the three groups after trimming for the three item types (J-only, E-only, and baseline) for the AN and VN categories combined (hereafter, for simplicity, *types* will be used to refer to J-only, E-only, and baseline differences, while *categories* will be used to refer to AN and VN).[6] Tables 4 and 5 show the results separated by AN and VN categories, respectively.

All data were analyzed with mixed effects modeling using the lme4 package in the R statistical software package (R Core Team, 2012). In line with precautionary suggestions proposed by Baayen (2008), all statistical models for RTs were run

https://doi.org/10.1017/S0142716414000113 Published online by Cambridge University Press

Table 5. *Mean log RTs (standard deviations) and ERs for VN items*

	RT (*SD*)			Priming Effect		ER (%)		
Group	J-Only	E-Only	Base	J-Only	E-Only	J-Only	E-Only	Base
NS	792 (240)	788 (254)	820 (263)	28	32	1.3	0.7	0.9
HNNS	869 (241)	853 (249)	873 (252)	4	20	0.9	1.5	0.8
LNNS	970 (339)	966 (367)	975 (343)	26	17	1.9	2.5	2.1

Note: RTs, Response times; ER, error rate; VN, verb–noun; NS, native speaker; HNNS, higher nonnative speaker; LNNS, lower nonnative speaker.

using the data that was trimmed after log transformation as well as untrimmed data and data trimmed prior to log transformation. In each case, however, the data trimmed after log transformation resulted in the best fitting models, so this data was used throughout the analyses. All mixed models, for both the RT analyses and the ER analyses, included person (the individual participants) and item (the individual items) as crossed random effects. Furthermore, each model used by person random slopes for trial (i.e., the order in which the items appeared on the participant's LDT) and length (the total number of letters in an item) to adjust for individual variations for these two factors (see Baayen, 2008). By-person random slopes allow the slope lines to vary for each participant to accommodate individual differences of the effects of particular variables. For example, some participants showed a strong tendency for faster RTs as the test went on, while others were more consistent throughout the test. The by-person random slopes allowed us to accommodate this without imposing a "one slope fits all" approach. The decision to include these additional random effects was based initially on a visual inspection of the data plots and later on model comparisons including and excluding these factors as random slope effects. As for fixed effects, all models included trial, length, and (log adjusted) lexical frequencies for both the first word (i.e., the adjective or verb) and the second word (the noun), and item type (J-only, E-only, or baseline). Category (AN or VN) was only included in the models that analyzed AN and VN responses combined (see below).

Of primary importance for the present study was performance within each proficiency group in respect to the three item types and within the two-word class categories. For this analysis, two sets of three models were run for each proficiency group independently (one set for RT data and one set for ER data). The first model (for both sets) included all items regardless of their status as AN or VN. The second model used only the data from the AN items and the third model used data from only the VN items. The results of the RT models are shown in Tables 6–8 while the results of the ER models are shown in Tables 9–11. As can be seen in the tables, none of the differences was statistically significant for either the RTs or the ERs for either NNS group. By contrast, the NS group did demonstrate significantly

https://doi.org/10.1017/S0142716414000113 Published online by Cambridge University Press

Table 6. *Results of mixed model comparing native speaker response times for J-only, E-only, and baseline items*

	All Items		AN Items		VN Items	
Fixed Effect	*t*	*p*	*t*	*p*	*t*	*p*
Intercept	86.48		59.28		72.1	
Trial	−4.33	<.0001	−4.63	<.0001	−3.54	.0004
Length	5.71	<.0001	4.65	<.0001	4.96	<.0001
Word 1 frequency	−5.93	<.0001	−4.03	<.0001	−4.34	<.0001
Word 2 frequency	−1.24	.2150	−1.08	.2802	−0.43	.6672
Baseline vs. English	−3.90	.0001	−2.21	.0271	−3.29	.0010
Baseline vs. Japanese	−1.97	.0489	−0.26	.7949	−2.72	.0066
English vs. Japanese	1.68	.0930	1.67	.0950	0.49	.6242
Adjective vs. verb	2.34	.0193	—	—	—	—

Note: AN, Adjective–noun; VN, verb–noun; Word 1 frequency, log adjusted frequency of the adjective or verb; Word 2 frequency, log adjusted frequency of the noun.

Table 7. *Results of mixed model comparing higher nonnative speaker response times for J-only, E-only, and baseline items*

	All Items		AN Items		VN Items	
Fixed Effect	*t*	*p*	*t*	*p*	*t*	*p*
Intercept	76.00		52.80		57.96	
Trial	−5.63	<.0001	−5.37	<.0001	−5.45	<.0001
Length	7.89	<.0001	5.98	<.0001	4.87	<.0001
Word 1 frequency	−5.91	<.0001	−3.65	.0003	−4.69	<.0001
Word 2 frequency	−2.97	.0030	−1.90	.0575	−2.03	.0425
Baseline vs. English	−1.00	.3174	−0.42	.6745	−1.03	.3031
Baseline vs. Japanese	0.73	.4654	0.93	.3525	0.01	.9920
English vs. Japanese	1.50	.1337	1.16	.2462	0.90	.3682
Adjective vs. verb	−0.67	.5029	—	—	—	—

Note: AN, Adjective–noun; VN, verb–noun; Word 1 frequency, log adjusted frequency of the adjective or verb; Word 2 frequency, log adjusted frequency of the noun.

faster response times for the E-only items over the baseline items. This result is not surprising. At the same time, however, the NS also demonstrated the curious tendency to respond significantly faster to the J-only items over the baseline items, with this difference seeming to be attributable primarily to accelerated RTs to the VN J-only items (again when compared to the baseline items, see Tables 6–8). This was an unexpected result, given that the J-only VN were largely nonexistent in the COCA. We now turn to a more detailed discussion of the findings.

https://doi.org/10.1017/S0142716414000113 Published online by Cambridge University Press

Table 8. *Results of mixed model comparing lower nonnative speaker response times for J-only, E-only, and baseline items*

	All Items		AN Items		VN Items	
Fixed Effect	*t*	*p*	*t*	*p*	*t*	*p*
Intercept	69.58		47.45		55.07	
Trial	−6.31	<.0001	−5.47	<.0001	−6.74	<.0001
Length	9.56	<.0001	6.81	<.0001	7.71	<.0001
Word 1 frequency	−5.80	<.0001	−3.53	.0004	−4.45	<.0001
Word 2 frequency	−3.96	<.0001	−2.96	.0031	−2.42	.0156
Baseline vs. English	−1.74	.0819	−1.12	.2628	−1.35	.1771
Baseline vs. Japanese	0.09	.9283	−0.05	.9601	0.13	.8966
English vs. Japanese	1.58	.1188	0.94	.3473	1.28	.2007
Adjective vs. verb	−1.56	.1188	—	—	—	—

Note: AN, Adjective–noun; VN, verb–noun; Word 1 frequency, log adjusted frequency of the adjective or verb; Word 2 frequency, log adjusted frequency of the noun.

Table 9. *Results of mixed model comparing native speaker error rates for J-only, E-only, and baseline items*

	All Items		AN Items		VN Items	
Fixed Effect	*t*	*p*	*t*	*p*	*t*	*p*
Intercept	38.65		23.45		32.79	
Trial	1.86	.0629	1.52	.1286	1.17	.2421
Length	−1.35	.1770	−1.07	.2847	−0.97	.3321
Word 1 frequency	0.18	.8572	1.80	.0720	−1.32	.1869
Word 2 frequency	−0.77	.4413	−0.86	.3899	−0.59	.5552
Baseline vs. English	−0.84	.4009	−0.99	.3223	−0.33	.7471
Baseline vs. Japanese	−0.52	.6031	−1.36	.1739	1.02	.3078
English vs. Japanese	0.22	.8259	−0.62	.5353	1.17	.2342
Adjective vs. verb	−0.73	.4654	—	—	—	—

Note: AN, Adjective–noun; VN, verb–noun; Word 1 frequency, log adjusted frequency of the adjective or verb; Word 2 frequency, log adjusted frequency of the noun.

DISCUSSION

Processing of L2 collocations

This study was undertaken with a primary goal of determining whether there would be activation of acceptable L1 collocational patterns when processing word combinations entirely in the L2. Specifically, we sought to test whether the lemma-based explanation for L2 collocational processing we developed drawing on the

https://doi.org/10.1017/S0142716414000113 Published online by Cambridge University Press

Table 10. *Results of mixed model comparing higher nonnative speaker error rates for J-only, E-only, and baseline items*

	All Items		AN Items		VN Items	
Fixed Effect	*t*	*p*	*t*	*p*	*t*	*p*
Intercept	28.70		17.30		26.56	
Trial	−0.19	.8493	−0.53	.5962	0.21	.8337
Length	0.85	.3954	0.77	.4414	0.46	.6456
Word 1 frequency	−2.63	.0086	−2.85	.0044	−0.95	.3422
Word 2 frequency	−1.04	.2984	−0.58	.5620	−0.61	.5419
Baseline vs. English	0.69	.4902	0.22	.8259	1.20	.2302
Baseline vs. Japanese	1.39	.1646	1.42	.1557	0.14	.8887
English vs. Japanese	0.61	.5419	1.02	.3078	−0.90	.3682
Adjective vs. verb	−2.86	.0043	—	—	—	—

Note: AN, Adjective–noun; VN, verb–noun; Word 1 frequency, log adjusted frequency of the adjective or verb; Word 2 frequency, log adjusted frequency of the noun.

Table 11. *Results of mixed model comparing low nonnative speaker error rates for J-only, E-only, and baseline items*

	All Items		AN Items		VN Items	
Fixed Effect	*t*	*p*	*t*	*p*	*t*	*p*
Intercept	32.33		23.09		22.80	
Trial	0.24	.8103	−0.41	.6818	1.09	.2758
Length	1.34	.1803	0.67	.5029	1.65	.0991
Word 1 frequency	−1.57	.1165	−0.70	.4840	−1.22	.2226
Word 2 frequency	−4.18	<.0001	−2.44	.0148	−3.53	.0004
Baseline vs. English	−0.34	.7339	−1.06	.2892	0.54	.5892
Baseline vs. Japanese	−1.38	.1676	−1.73	.0837	−0.31	.7566
English vs. Japanese	−0.90	.3682	−0.57	.5687	−0.75	.4533
Adjective vs. verb	1.22	.2225	—	—	—	—

Note: AN, Adjective–noun; VN, verb–noun; Word 1 frequency, log adjusted frequency of the adjective or verb; Word 2 frequency, log adjusted frequency of the noun.

work of Jiang (2000) provided a plausible explanation for the accelerated processing of congruent collocations in the L2 seen in previous studies (Yamashita & Jiang, 2010; Wolter & Gyllstad, 2011, 2013). Based on the findings, we are forced to conclude that the apparent advantage afforded to L2 learners in respect to processing congruent L2 collocations may not be due to activation of L1-appropriate links. In other words, there is no evidence emerging from this study to support the lemma-based explanation for L2 collocational development proposed

https://doi.org/10.1017/S0142716414000113 Published online by Cambridge University Press

above. Another explanation, one that was raised in the introductory section, is the AoA or OoA effect (Carroll & White, 1973). To recap, the basic idea here is that congruent collocations tend to be acquired earlier than incongruent collocations and this in turn leads to faster processing for congruent over incongruent collocations. Applying these theories to the research literature on congruent versus incongruent collocational processing provides us with an account for why congruent collocations tend to be processed faster than incongruent collocations, and we believe that AoA and OoA provide plausible explanations that should be further explored empirically.

Another finding to emerge from this study that needs further elaboration was that E-only items also did not elicit significantly faster RTs for either NNS group. These items were, in essence, what have been referred to in other studies as incongruent collocations because they had no translation equivalent in the NNSs' L1. However, the results of this study are not consistent with those of previous studies (Wolter & Gyllstad, 2011, 2013). Both of these studies reported significantly faster recognition among advanced L1 Swedish learners for incongruent collocations over baseline items. The question here is how to explain this difference. We feel that these conflicting findings might be at least partially attributable to fundamental differences in the L1s and the L2 learning experiences of the participants involved in the respective studies (i.e., Swedish vs. Japanese). As is well known, English and Swedish are closely related languages with shared canonical word order and (important for lexical decision) more or less shared orthographies. Furthermore, Swedish learners will generally have much greater access to English input in their daily lives, which can only serve to further strengthen and solidify their L2 language systems (and ultimately their proficiency). These conditions are clearly quite distinct from L1 Japanese learners, whose language has much less in common with English, including vastly different orthographies, and whose L2 input is likely to be considerably less than an L1 Swedish learner's.

These important differences notwithstanding, we feel that more than anything the E-only RT results provide a measure of confirmation for the difficulty involved in effectively processing (and perhaps fully acquiring) L2 collocations that have no congruent form in the L1. Although Wolter and Gyllstad (2011, 2013) reported significantly faster RTs to incongruent collocations over baseline items, it is worth noting that the incongruent RTs were also significantly slower than they were for congruent collocations in both studies. Furthermore, in their first study Wolter and Gyllstad (2011) reported two important findings regarding incongruent collocations. The first was that there was a discrepancy between the $F1$ (with participants as random variables) and $F2$ analyses in terms of significance. In particular, although there was a significant difference for the $F1$ analysis between RTs in the incongruent and baseline condition ($p < .001$), this was not replicated for the $F2$ analysis ($p = .106$). Based on this, Wolter and Gyllstad concluded that the priming demonstrated across the incongruent condition for the participants was likely not generalizable to all the items in the incongruent condition. In other words, they suggested, priming was affected by more varied recognition of the collocations in the incongruent condition than was seen in the congruent condition. The second finding was in respect to results Wolter and Gyllstad reported on a test of receptive

https://doi.org/10.1017/S0142716414000113 Published online by Cambridge University Press

collocational knowledge (COLLMATCH; Gyllstad, 2007). These results showed significantly better recognition of legitimate congruent collocations over legitimate incongruent collocations ($p < .001$). In other words, the participants simply knew more congruent than incongruent collocations, even though the collocational frequency was matched across the conditions. When these findings are considered in respect to the results of the current study (and those obtained by Yamashita & Jiang, 2010), it leads us to the conclusion that incongruent collocations really do continue to pose processing challenges for L2 learners, even at higher levels of proficiency.

The last aspect of the L2 participants' results we would like to consider is in respect to trends in the data that indicate the L2 participants in both proficiency groups may have employed a fundamentally different approach to the task than the NSs in this study. Specifically there is evidence that indicates that NNSs processed items on the LDT less holistically than the NSs. This is because, for the two NNS groups, the frequency of the second word in the items had a significant effect on RTs (Tables 7 and 8) whereas it did not have such an effect for the NS group (Table 6). This implies that the NNSs may have been more serial and disjointed in their approach on the LDT than the NSs, which in turn indicates that some of the NNSs may have managed to complete the task without actually accessing collocational knowledge (or at least presumably not to the degree that collocational knowledge was accessed by the NSs). If this was the case, then it suggests that there may be cause to replicate the current study with some modifications to the online task. Specifically, it may be of value to gauge RTs (and ERs) on a task that more strictly requires access to collocational knowledge before we definitively dismiss the idea that L1 collocational connections are not activated during L2 collocational processing.

Processing collocations in the L1

Although it was not our initial intention to treat the results of the NS group in this study as anything other than control data, we would be remiss if we failed to address the curious findings we observed for the NS RTs to the J-only collocations. To recap, the NS group unexpectedly demonstrated significantly faster RTs to the J-only items over the baseline items, even though these J-only collocations were almost completely absent in the COCA. Further analyses revealed that this was due to accelerated RTs to the VN J-only items rather than the AN J-only items. The obvious question that arises here is what could account for this behavior.

One possible explanation is that there is a fundamental difference in how AN and VN combinations are processed in one's L1. In brief, verbs prime not only nouns that are found in the L1 speaker's input but also other nouns that could conceptually be associated with the meaning represented by a particular verb. This would explain why J-only VN combinations elicited accelerated RTs even though randomly determined VN combinations (consisting entirely of the same words found in the collocational items) did not. It is our sense that although the VN J-only items on the test were truly not natural collocations in English (see Appendixes A and B), that they are natural in another language means that they

https://doi.org/10.1017/S0142716414000113 Published online by Cambridge University Press

represent some sort of conceivable conceptual connection between the meaning of the verb and the meaning of the noun that could manifest itself at more of a universal level that is not specific to a particular language. In this respect, it could be that the activation we see between verbs and nouns occurs more on a semantic level than on input/frequency-based level.

However, for AN combinations the situation might be quite different. Again, it is our estimation that the AN J-only items did not represent acceptable collocations in English. However, it is not hard to decipher what many of these mean, even in the absence of knowledge regarding the Japanese language. In this respect, it cannot be said that there is no clear conceptual connection between many of the adjectives and the nouns in the J-only condition. Furthermore, it should be noted that when presented with the task of judging whether random AN combinations represented a "common expression in English," as was tested in the study by Wolter and Gyllstad (2013), both NSs and advanced NNSs (L1 Swedish) showed a tendency to accept combinations that did not occur in the 450-million word COCA. On average, NSs judged 27.8% of these combinations as commonly occurring while the figure for the NNSs was 28.8%. Thus, research has indicated that people will consciously judge nonattested AN collocations as commonly occurring. There is practical value in being able to accept novelty of expression for skillful language processing. However, conscious judgment is not the same as processing; although we believe even non-Japanese speaking people could see the conceptual connection between the adjectives and nouns in many of the J-only AN items, these items, in contrast to the VN items, were not processed any more rapidly than the baseline AN items. Therefore, we would like to tentatively propose the possibility that there may be a difference in the language system for processing AN versus VN items. Again, for VN combinations it may be that there is essentially semantic activation that is less conditioned by previous experience with the combinations of the particular words involved, while for AN combinations it may be much more experience driven.

From a theoretical perspective, there is some justification for this view. This explanation has to do with the concept of valency (Tesnière, 1959) or, as it has been called by Chomsky (1965), subcategorization. The basic idea behind these theories is that verbal argument structures play a pivotal role in assigning function to the noun phrases in the clause (see, e.g., Jackendoff, 1983; Pinker, 1989). Furthermore, the argument structure of a verb becomes learned and stored as part of the word's entry in the mental lexicon. What emerges, then, is a situation in which the information stored in the lexical entry at the semantic level has ramifications for the processing of subsequent language at a syntactic level. If it can be assumed that this state of affairs for verbs in the mental lexicon affects language processing, then one can envision a situation in which skillful language users presented with a particular transitive verb will subconsciously prime nouns that could conceptually be connected to the verb, even if these are not part of his or her L1 language experience.

What about adjectives, though? As noted earlier, we believe that there were many clear conceptual connections between the J-only AN combinations, and we suspect that the meaning of many of these expressions could correctly be recognized by people with no knowledge of Japanese. At the same time, however, these items

https://doi.org/10.1017/S0142716414000113 Published online by Cambridge University Press

did not invoke the sort of accelerated processing we witnessed with the J-only VN items. The obvious question that arises here is why not. The answer again lies in the concept of valency/subcategorization. Whereas verbs subcategorize noun phrases, adjectives are typically subsumed within noun phrases, and as such, they do not have the same sort of syntactic sway that verbs have under this system. Simply put, verbs subcategorize noun phrases, but adjectives subcategorize nothing. If this descriptive observation about language systems ends up having a bearing on how language is internalized psychologically, then it might lead to the previously proposed situation in which adjectives come to prime only nouns that they commonly co-occur within one's language input while verbs are more versatile in the nouns they can come to prime.

What we are describing is something of a hybrid system that incorporates ideas from two perspectives, which have been ideologically opposed to one another, in an effort to explain the empirical results obtained in this study. The idea of subcategorization and argument structure clearly have their roots in Chomksyan linguist theories, while the importance of frequency in input (and input in general) is a centerpiece of the various forms of usage-based grammars. Historically speaking, the biggest point of contention between these two schools of thought has been in respect to the existence of innate linguistics knowledge. We can see no reason, however, why the concept of subcategorization cannot be successfully applied to theories of language that assume no innate linguistic knowledge, such as that found in the various usage-based grammars. Usage-based grammars already talk about the commonality for some lexical items to have dual status as both a (grammatical) construction and a lexical item. Langacker (2000) provides us with the example of the ditransitive verb *send* and the patterning it takes in the expressions that take the form of *send [NP] [NP]*. He argues that it is unproductive to attempt to rule it as either representing a generic case of a *V [NP] [NP]* construction or as a single lexical item consisting specifically of *send [NP] [NP]*. Instead, he argues, it is both simultaneously. Applying this thinking to the current study gives us a way of reconciling the two views. It is possible that the tendency of verbs to subcategorize noun phrases allows them to becoming generalized constructions that can more readily be abstracted from examples in the language input. These generalized constructions could then be applied to subsequent *V [NP]* patterns, even when these have not been a part of the language user's input. In contrast, adjectives would tend to retain more of a lexical item status; in this sense, their activation would continue to be linked to actual encounters in the language user's input.

These explanations are speculative, but they are driven by an effort to explain some highly unanticipated findings in our NS data. The evidence that has emerged from this study supports the claim that verb–(object)–noun combinations that do not exist in a speaker's L1, but are plausible in another language, may offer processing advantages over randomly generated verb–(object)–noun combinations. No such advantage was seen for the AN items, despite their acceptability in Japanese. Nonetheless, it will first of all be necessary to replicate these findings to see if they hold with other groups of NSs (and if so, perhaps also with translations of incongruent collocations drawn from other languages). Until this happens we

https://doi.org/10.1017/S0142716414000113 Published online by Cambridge University Press

cannot be sure that the findings in this study represent a truly generalizable trend in L1 collocational processing rather than something that is simply anomalous and confined to the present study.

Before pursuing such inquiries, however, it would also be worthwhile to more closely investigate one more factor that distinguished the AN items from the VN items in the current study, which is that the AN items tended to be, on the whole, more grammatically correct than the VN items because the VN items lacked determiners that would sometimes be present in natural language use. This was an adjustment that was incorporated in order to accommodate both types of items in the format of the dual LDT. Nonetheless, it is not entirely clear if this affected the results or not. In other words, although it still does nothing to explain the NSs' tendency to respond more quickly to the J-only VN items over the baseline VN items, we cannot be sure that this difference did not have some bearing on the NSs' different response patterns for these two-word class categories. In brief, it may be useful to replicate the current study using only VN items that can still be considered grammatically correct without inclusion of a determiner (e.g., through greater use of the plural forms of the nouns as in *draw conclusions* instead of *draw conclusion*).

Conclusion

The main aim of this study was to test a key assumption of a description of L2 collocational development that was based on Jiang's (2000) model of lexical acquisition. The description of collocational development was proposed by Wolter and Gyllstad (2011) in an effort to account for previous research findings indicating a processing advantage for L2 collocations that had a congruent L1 form (i.e., collocations that were felicitous in both the L1 and the L2) over those that had no congruent form (i.e., collocations that were acceptable in the L2 but not in the L1; Wolter & Gyllstad, 2011, 2013; Yamashita & Jiang, 2010). More specifically, the description assumed the possibility that L1 collocational information may be copied into the L2 lexical entry at the lemma level (see Figure 1), and that this in turn may cause the activation of L1-appropriate collocations even when processing collocations entirely in the L2. To disentangle the potential influence of L2 knowledge in the task, translations of collocations that were felicitous in the L1 but not the L2 (e.g., *far eye*, which is acceptable in Japanese but not English) were used in a study employing a double LDT. The results revealed no accelerated activation of the L1-appropriate collocations, leading us to conclude that the description may not provide an accurate account of L2 collocational development. As an alternative, we suggested the possibility that AoA/OoA effects (Carroll & White, 1973) may provide a better explanation for the congruency effect observed in the previous studies investigating L2 collocational processing. However, this view will need to be empirically tested.

Nonetheless, we still feel that there may be cause to revisit the lemma-based description, and the associated issue of activation of L1-appropriate links for L2 collocations, in future studies. Specifically, we believe that modifications to the

https://doi.org/10.1017/S0142716414000113 Published online by Cambridge University Press

experimental design that require more explicit focus on collocational knowledge per se (over individual word knowledge or even knowledge of language forms) may produce different and more revealing results. In short, we feel that it may be too early to reject the lemma-based description entirely based solely on the results of the current study.

Finally, we were surprised to find that NSs demonstrated accelerated RTs to translations of verb–(object) noun collocations that were acceptable in Japanese but not in English. No such pattern was found for translations of Japanese AN collocations. Based on this, we put forth some arguments that draw both on sub-categorization, a key concept in Chomskyan linguistics, and (lexical) abstraction that is an idea found mainly in usage-based grammars.

APPENDIX A

Adjective–noun items on lexical decision task

J-Only Collocations	E-Only Collocations	Baseline	
Active personality	Bad debt	Active hope	Low parents
Bitter win	Bare feet	Bad gift	Lucky learner
Cheap pay	Big advantage	Bare content	Narrow judgment
Dark society	Brief moment	Big moment	Open health
Delicious season	Broad daylight	Bitter bird	Poor rise
Far eye	Busy road	Brief heart	Pretty advantage
Frank question	Careful attention	Broad luck	Proud effect
High effect	Clean break	Busy teeth	Quick rain
Hot battle	Coming months	Careful branch	Quiet rate
Light content	Early bird	Cheap feet	Severe energy
Narrow thought	Easy decision	Clean market	Shallow season
Pretty animal	False teeth	Coming look	Sharp pay
Quiet war	Flat rate	Dark guess	Short personality
Severe home	Free speech	Delicious speed	Silent arms
Shallow injury	Full understanding	Early voice	Slim decision
Silent force	Hard luck	Easy force	Slow eye
Slim branch	Heavy traffic	False months	Soft road
Soft atmosphere	Low speed	Far understanding	Straight daylight
Straight heart	Lucky guess	Flat atmosphere	Strong question
Strong hope	Open arms	Frank traffic	Sweet win
Sweet judgment	Poor health	Free injury	Tall notice
Tall danger	Proud parents	Full society	Warm thought
Warm gift	Quick look	Hard home	Weak war
Weak rain	Sharp rise	Heavy battle	Wet attention
Wide market	Short notice	High break	Wide danger
Yellow voice	Slow learner	Hot speech	Yellow weather
Young energy	Wet weather	Light animal	Young debt

https://doi.org/10.1017/S0142716414000113 Published online by Cambridge University Press

APPENDIX B

Verb–noun items on lexical decision task

J-Only Collocations	E-Only Collocations	Baseline	
Add harm	Answer phone	Add feelings	Forgive role
Bend fact	Attach importance	Answer shoulder	Give legs
Borrow toilet	Bite bullet	Attach distance	Have harm
Break health	Bow head	Bend patient	Hit vote
Bring legs	Cast vote	Bite importance	Kill attack
Buy anger	Catch breath	Borrow rain	Lose visit
Crash time	Draw conclusion	Bow breath	Make tears
Cut telephone	Eat soup	Break ears	Meet contest
Dance heart	Enter contest	Bring condition	Pass marriage
Decide mind	Exercise power	Buy eyes	Pay war
Deliver eyes	Fight war	Cast shower	Play voice
Do baseball	Fly flag	Catch health	Pour chance
Doubt ears	Lose game	Crash mistakes	Push judgment
Drink tears	Make mistakes	Cut anger	Put doctor
Fill condition	Pass judgment	Dance game	Raise soup
Forgive marriage	Pay visit	Decide breast	Receive party
Give influence	Play role	Deliver head	Save baseball
Have shoulder	Push luck	Do heart	See telephone
Hit breast	Save face	Doubt toilet	Spread bullet
Kill voice	See doctor	Draw flag	Stand score
Meet rain	Spread word	Drink luck	Stretch information
Pour love	Stand chance	Eat effect	Switch love
Put distance	Take shower	Enter conclusion	Take influence
Raise effect	Throw party	Exercise time	Throw face
Receive attack	Treat patient	Fight word	Treat mind
Stretch score	Volunteer information	Fill dog	Volunteer phone
Switch feelings	Walk dog	Fly fact	Walk power

ACKNOWLEDGMENTS

We thank Naohiro Takizawa for his generous help in the analysis of the Japanese corpus.

NOTES

1. Note that we are using the term *priming* here to refer to priming in a technical, empirical sense. In other words, we are describing the phenomenon of recognizing a word that has been preceded by a related prime word more quickly than an unrelated prime word. This is different than the sense in which the term is used by Hoey (2005), who also refers to collocational priming but more in a theoretical than technical sense.
2. Most individual words in the stimuli were within the most frequent 3,000-word level in the JACET List of 8000 Basic Words (JACET Committee of Basic Words Revision,

https://doi.org/10.1017/S0142716414000113 Published online by Cambridge University Press

2003). This is the level for high school textbooks in the Japanese education system; therefore, the words were assumed to be known to the participants. Some words from the lower frequency bands were also assumed to be known because they were either derivatives (e.g., *learner*, *coming*) or compounds (e.g., *daylight*) of the high frequency words. The words over the 3,000-word level included *frank* (5,574), *slim* (4,302), *coming* (3,997), *learner* (4,234), and *daylight* (4,388).

3. We used the Chunichi Shinbun (Chunichi Newspaper) Database (1999 to 2003).
4. Although raw frequencies were not used for decisions regarding inclusion of items, the mean raw frequency score was 388 occurrences in the corpus with a low frequency of 27 and a high frequency of 2,385 (again for *lucky guess* and *free speech* respectively).
5. As with the MI scores, the frequency scores were calculated slightly differently for the AN and VN scores. The AN frequencies were calculated by obtaining instances of contiguous co-occurrence in the COCA. For the VN frequencies, however, we used the sum of the number of contiguous co-occurrences and the number of co-occurrences of the verb and the noun with one (undefined) determiner separating the two. This allowed us to account for variations in the VN items, because some items tended to occur more frequently without an intermediary determiner (e.g., *save face*) while others (most items) tended to occur more commonly with a determiner (e.g., *catch breath*).
6. All statistical analyses involving RTs were conducted using log-transformed data. However, for ease of understanding, the tables show the base of the transformed data. These figures were computed by converting each remaining log transformed RT after trimming back into its base form and then calculating the average and standard deviations for these numbers collectively.

REFERENCES

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

Bahns, J. (1993). Lexical collocations: A contrastive view. *ELT Journal, 47*, 56–63.

Bahns, J., & Eldaw, M. (1993). Should we teach EFL students collocations? *System, 21*, 101–114.

Barfield, A., & Gyllstad, H. (2009). *Researching collocations in another language-multiple interpretations*. New York: Palgrave Macmillan.

Brysbaert, M., & Duyck, W. (2010). Is it time to leave behind the Revised Hierarchical Model of bilingual language processing after fifteen years of service? *Bilingualism: Language & Cognition, 13*, 359–371.

Brysbaert, M., Van Wijnendaele, I., & De Deyne, S. (2000). Age of acquisition effects in semantic tasks. *Acta Psychologica, 104*, 215–226.

Bybee, J. L. (2007). *Frequency of use and the organization of language*. Oxford: Oxford University Press.

Bybee, J., & Hopper, P. (2001). Introduction. In J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins.

Carroll, J. B., & White, M. N. (1973). Word frequency and age of acquisition as determiners of picture-naming latency. *Quarterly Journal of Experimental Psychology, 25*, 85–95.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chwilla, D. J., & Kolk, H. H. J. (2002). Three-step priming in lexical decision. *Memory & Cognition, 30*, 217–225.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*, 407–428.

Cowie, A. P. (1998). *Phraseology: Theory, analysis, and applications*. Oxford: Clarendon Press.

Davies, M. (2008). *The Corpus of Contemporary American English: 450 million words, 1990–present.* Retrieved from http://corpus.byu.edu/coca

https://doi.org/10.1017/S0142716414000113 Published online by Cambridge University Press

Durrant, P. & Doherty, A. (2010). Are high-frequency collocations psychologically real? Investigating the thesis of collocational priming. *Corpus Linguistics and Linguistic Theory, 6*, 125–311.

Ellis, A. W., & Lambon Ralph, M. A. (2000). Age of acquisition effects in adult lexical processing reflect loss of plasticity in maturing systems: Insights from connectionist networks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1103–1123. doi:10.1037/0278-7393.26.5.1103

Ellis, N., Frey, E., & Jalkanen, I. (2009). The psycholinguistic reality of collocation and semantic prosody. In U. Römer & R. Schulze (Eds.), *Exploring the lexis-grammar interface*. Amsterdam: John Benjamins.

Fan, M. (2009). An exploratory study of collocational use by ESL students: A task-based approach. *System, 37*, 110–123.

Firth, J. R. (1957). A synopsis of linguistic theory, 1930–1955. In F. R. Palmer (Ed.), *Selected papers of J.R. Firth*. London: Longmans.

Fischler, I. (1977). Semantic facilitation without association in a lexical decision task. *Memory & Cognition, 5*, 335–339.

Gabrys-Biskup, D. (1992). L1 influence on learners' renderings of English collocations: A Polish/German empirical study. In P. J. L. Arnaud & H. Bejoint (Eds.), *Vocabulary and applied linguistics* (pp. 85–93). New York: Macmillan.

Gass, S. (1988). Second language acquisition and linguistic theory: The role of language transfer. *Studies in Theoretical Psycholinguistics, 8*, 384–403.

Ghyselinck, M., Lewis, M. B., & Brysbaert, M. (2004). Age of acquisition and the cumulative-frequency hypothesis: A review of the literature and a new multi-task investigation. *Acta Psychologica, 115*, 43–67.

Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.

Granger, S. (1998). Prefabricated patterns in advanced EFL writing: Collocations and formulae. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis and application*. Oxford: Oxford University Press.

Gyllstad, H. (2007). *Testing English collocations: Developing receptive tests for use with advanced Swedish learners*. Unpublished doctoral dissertation, Lund University.

Hoey, M. (2005). *Lexical priming: A new theory of words and language*. New York: Routledge.

Howarth, P. (1998). Phraseology and second language proficiency. *Applied Linguistics, 19*, 24–44.

Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.

Izura, C., & Ellis, A. W. (2002). Age of acquisition effects in word recognition and production in first and second languages. *Psicológica, 23*, 245–281.

Izura, C., & Ellis, A. W. (2004). Age of acquisition effects in translation judgement tasks. *Journal of Memory and Language, 50*, 165–181.

Izura, C., Pérez, M. A., Agallou, E., Wright, V. C., Marín, J., Stadthagen-González, H., et al. (2011). Age/order of acquisition effects and the cumulative learning of foreign words: A word training study. *Journal of Memory and Language, 64*, 32–58.

JACET Committee of Basic Words Revision. (Ed.). (2003). *JACET list of 8000 basic words: JACET 8000*. Tokyo: Japan Association of College English Teachers.

Jackendoff, R. (1983). *Semantics and cognition*. Cambridge, MA: MIT Press.

Jiang, J. (2009). Designing pedagogic materials to improve awareness and productive use of L2 collocations. In A. Barfield & H. Gyllstad (Eds.), *Researching collocations in another language: Multiple interpretations* (pp. 99–113). Basingstoke: Palgrave Macmillan.

Jiang, N. (2000). Lexical representation and development in a second language. *Applied Linguistics, 21*, 47–77.

Jiang, N. (2004). Semantic transfer and its implications for vocabulary teaching in a second language. *Modern Language Journal, 88*, 416–432.

Jiang, N. (2012). *Conducting reaction time research*. New York: Routledge.

Jiang, N., Novokshanova, E., Masuda, K., & Wang, X. (2011). Morphological congruency and the acquisition of L2 morphemes. *Language Learning, 61*, 940–967.

Jones, L. L., & Estes, Z. (2012). Lexical priming: Associative, semantic, and thematic influences on word recognition. *Visual Word Recognition, 2*, 44–72.

Kasper, G. (1992). Pragmatic transfer. *Second Language Research, 8*, 203–231.

https://doi.org/10.1017/S0142716414000113 Published online by Cambridge University Press

Kellerman, E. (1979). Transfer and non-transfer: Where we are now? *Studies in Second Language Acquisition, 2*, 37–57.
Kirsner, K., Brown, H., Abrol, S., Chadha, N. K., & Sharma, N. K. (1980). Bilingualism and lexical representation. *Quarterly Journal of Experimental Psychology, 32*, 585–594.
Lambon Ralph, M. A., & Ehsan, S. (2006). Age of acquisition effects depend on the mapping between representations and the frequency of occurrence: Empirical and computational evidence. *Visual Cognition, 13*, 928–948.
Langacker, R. W. (1987). *Foundations of cognitive grammar: Theoretical prerequisites*. Stanford, CA: Stanford University Press.
Langacker, R. W. (1988). A usage-based model. In B. Rudzka-Ostyn (Ed.), *Topics in cognitive linguistics*. Amsterdam: John Benjamins.
Langacker, R. W. (2000). A dynamic usage-based model. In M. Barlow & S. Kemmer (Eds.), *Usage-based models of language*. Stanford, CA: CSLI Publications.
Lemhöfer, K., & Dijkstra, T. (2004). Recognizing cognates and interlingual homographs: Effects of code similarity in language-specific and generalized lexical decision. *Memory & Cognition, 32*, 533–550.
Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
Lucas, M. (2000). Semantic priming without association: A meta-analytic review. *Psychonomic Bulletin & Review, 7*, 618–630.
MacWhinney, B. (2000). Connectionism and language learning. In M. Barlow & S. Kemmer (Eds.), *Usage-based models of language* (pp. 121–149). Stanford, CA: CSLI Publications.
Manelis, L., & Tharp, D. A. (1977). The processing of affixed words. *Memory & Cognition, 5*, 690–695.
Marinis, T., Roberts, L., Felser, C., Clahsen, H. (2005). Gaps in second language sentence processing. *Studies in Second Language Acquisition, 27*, 53–78.
McDonald, S. A., & Shillcock, R. C. (2003a). Eye movements reveal the on-line computation of lexical probabilities during reading. *Psychological Science, 14*, 648–652.
McDonald, S. A., & Shillcock, R. C. (2003b). Low-level predictive inference in reading: The influence of transitional probability on eye movements. *Vision Research, 43*, 1735–1751. \
McDonough, K., & Trofimovich, P. (2009). *Using priming methods in second language research*. New York: Routledge.
McNamara, T. P. (2004). *Semantic priming: Perspectives from memory and word recognition*. New York: Psychology Press.
Meara, P., & Jones, G. (1990). *Eurocentres vocabulary size tests 10K*. Zurich: Eurocentres Learning Service.
Meyer, D. E., & Schvaneveldt, R. W. (1971) Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90*, 227–234. doi:10.1037/h0031564
Monaghan, J., & Ellis, A. W. (2002). What exactly interacts with spelling–sound consistency in word naming? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 183.
Myerson, J., Hale, S., Chen, J., & Lawrence, B. (1997). General lexical slowing and the semantic priming effect: The roles of age and ability. *Acta Psychologica, 96*, 83–101.
Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics, 24*, 223–242.
Papadopoulou, D., & Clahsen, H. (2003). Parsing strategies in L1 and L2 sentence processing: A study of relative clause attachment in Greek. *Studies in Second Language Acquisition, 25*, 501–528.
Perea, M., & Rosa, E. (2002). Does the proportion of associatively related pairs modulate the associative priming effect at very brief stimulus-onset asynchronies? *Acta Psychologica, 110*, 103–124.
Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: MIT Press.
R Core Team. (2012). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 nonwords: The ARC nonword database. *The Quarterly Journal of Experimental Psychology, 55*, 1339–1362.

https://doi.org/10.1017/S0142716414000113 Published online by Cambridge University Press

Sunderman, G., & Kroll, J. F. (2006). First language activation during second language lexical processing. *Studies in Second Language Acquisition, 28*, 387–422.
Schmitt, N., & Carter, R. (2004). Formulaic sequences in action: An introduction. In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing, and use*. Philadelphia, PA: John Benjamins.
Schoonbaert, S., Duyck, W., Brysbaert, M., & Hartsuiker, R. J. (2009). Semantic and translation priming from a first language to a second and back: Making sense of the findings. *Memory & Cognition, 37*, 569–586.
Sinclair, J. (1991). *Corpus concordance collocation*. Oxford: Oxford University Press.
Siyanova-Chanturia, A., Conklin, K. & van Heuven, W. J. B. (2011). Seeing a phrase "time and again" matters: The role of phrasal frequency in the processing of multiword sequences. *Journal of Experimental Psychology, 37*, 776–784.
Tesnière, L. (1959). *Eléments de syntaxe structural*. Paris: Klincksieck.
Tomasello, M. (2000). First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics, 11*, 61–82.
Tomasello, M. (2003). *Constructing a language: A usage-based approach to language acquisition*. Cambridge, MA: Harvard University Press.
Van Loon-Vervoom, W. A., Van Ham, P., & Van der Koppen, M. (1988). The importance of age of acquisition for imageability in word processing. In M. Dennis, J. Engelkamp, & J. T. E. Richardson (Eds.), *Cognitive and neuropsychological approaches to mental imagery* (pp. 99–107). Dordrecht: Martinus Nijhoff.
Wang, M., Koda, K., & Perfetti, C. A. (2003). Alphabetic and nonalphabetic L1 effects in English word identification: A comparison of Korean and Chinese English L2 learners. *Cognition, 87*, 129–149.
Wolter, B., & Gyllstad, H. (2011). Collocational links in the L2 mental lexicon and the influence of L1 intralexical knowledge. *Applied Linguistics, 32*, 430–449.
Wolter, B., & Gyllstad, H. (2013). Frequency of input and L2 collocational processing. *Studies in Second Language Acquisition, 35*, 451–482. doi:10.1017/S0272263113000107
Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.
Wray, A. (2008). *Formulaic language: Pushing the boundaries*. Oxford: Oxford University Press.
Yamashita, J., & Jiang, N. (2010). L1 influence on the acquisition of L2 collocations: Japanese ESL users and EFL learners acquiring English collocations. *TESOL Quarterly, 44*, 647–668.

https://doi.org/10.1017/S0142716414000113 Published online by Cambridge University Press