Studies in Second Language Acquisition, 2016, **38**, 403–443.
doi:10.1017/S0272263115000492

FOUND IN TRANSLATION

The Influence of the L1 on the Reading of Idioms in a L2

Gareth Carrol and Kathy Conklin
University of Nottingham

Henrik Gyllstad
Lund University

Formulaic language represents a challenge to even the most proficient of language learners. Evidence is mixed as to whether native and nonnative speakers process it in a fundamentally different way, whether exposure can lead to more nativelike processing for nonnatives, and how L1 knowledge is used to aid comprehension. In this study we investigated how advanced nonnative speakers process idioms encountered in their L2. We used eye-tracking to see whether a highly proficient group of L1 Swedes showed any evidence of a formulaic processing advantage for English idioms. We also compared translations of Swedish idioms and congruent idioms (items that exist in both languages) to see how L1 knowledge is utilized during online processing. Results support the view that L1 knowledge is automatically used from the earliest stages of processing, regardless of whether sequences are congruent, and that exposure and advanced proficiency can lead to nativelike formulaic processing in the L2.

Correspondence concerning this article should be addressed to Gareth Carrol, School of English, University of Nottingham, Nottingham, NG7 2RD, United Kingdom. E-mail: gareth.carrol@nottingham.ac.uk

INTRODUCTION

Alongside the acquisition of sufficient vocabulary and grammatical competence, "nativelike" proficiency in a language requires mastery of the vast array of word strings and conventionalized sequences that characterize native speaker interaction. This broad category of lexical knowledge is considered under the banner of *formulaic language*. Such multiword combinations may be at least as numerous as the amount of single words in English (Jackendoff, 1995), possibly numbering into the hundreds of thousands (Pawley & Syder, 1983). Crucially, they present an ongoing challenge to nonnative speakers, even at advanced levels of proficiency (Barfield & Gyllstad, 2009; Laufer & Waldman, 2011). Increasing attention has been paid to formulaic language in language learning, and to how such combinations are represented in the mental lexicon for both native and nonnative speakers.

Idioms—noncompositional sequences of words that denote a specific figurative meaning—arguably pose the greatest degree of difficulty for nonnative speakers. Idioms often behave like single words in performing a referential or ideational function (Boers & Lindstromberg, 2012), but their difficulty for language learners comes from the fact that they are often opaque, and their meaning difficult to infer without some prior knowledge. Their importance to the study of formulaic language is underlined by the claim from Titone, Columbus, Whitford, Mercier, and Libben (2015) that "idioms optimally represent the larger class of MWEs [multiword expressions] as they vary along all linguistic dimensions relevant to MWEs generally, including familiarity, literal plausibility, semantic decomposability, and other linguistic attributes" (p. 173). Idioms are therefore best seen as existing along various continua of formulaicity, including one of figurativeness, with fully opaque idioms at one end and transparent but frequently occurring phrases at the other. It is not always possible to strictly demarcate idioms from other types of formulaic sequence such as semitransparent collocations or phrasal verbs, but a key defining feature is that idioms are lexically frozen and otherwise fixed in highly conventionalized ways.

As we discuss in detail subsequently, evidence is mixed as to how nonnative speakers process, comprehend, and produce idioms, and how L1 knowledge is utilized to support their use in communication. There is still something of a research gap in terms of constructing a detailed model of how idioms and other types of formulaic language are represented and processed by L2 speakers. To help address this, in the present study we investigated how nonnative speakers process idioms that they encounter in their L2. Specifically, we presented advanced learners of English with idioms in three categories—L2-only idioms, translations of L1-only idioms, and idioms that consist of the same

https://doi.org/10.1017/S0272263115000492 Published online by Cambridge University Press

combination of words and the same phrase-level meaning in both languages—to see to what extent L1 knowledge was utilized and how this interacted with L2 formulaic competence. To begin, we review two principle strands of previous research: the psycholinguistic literature on the processing of idioms in the L1 and L2, and those studies that have investigated formulaic transfer from the L1 in nonnative speakers.

Formulaic Processing in the L1 and L2: Different Strokes for Different Folks

The Formulaic Processing Advantage in Native Speakers. It is well established that idioms and other types of formulaic language are processed more quickly than "novel language" by native speakers when other factors like length and single-word frequency are controlled for. It is important to note that novel language need not be entirely new in the sense of having never been heard before. Rather, novel sequences are considered to be nonrecurrent combinations that do not show any significant degree of cohesion or fixedness, whereas known combinations are highly frequent, are highly cohesive, and/or have a single phrasal meaning. It has been demonstrated, using a range of methodologies, that idioms are processed differently from novel language (Cacciari & Tabossi, 1988; Conklin & Schmitt, 2008; Libben & Titone, 2008; McGlone, Glucksberg, & Cacciari, 1994; Rommers, Dijkstra, & Bastiaansen, 2013; Schweigert, 1986; Swinney & Cutler, 1979). The same is true of other types of formulaic sequence that can be seen as at least partially figurative, such as phrasal verbs (Blais & Gonnerman, 2013; Matlock & Heredia, 2002) and irreversible binomials such as *hit and run* (Arcara et al., 2012). This difference is most often apparent in the speed of processing, with faster processing often inferred to be an indicator of "whole-form" storage at some level of representation. For example, Swinney and Cutler (1979) show that an idiom like *break the ice* is judged to be a meaningful phrase more quickly than a control phrase like *break the cup*. Although certain recent research (e.g., Cutter, Drieghe, & Liversedge, 2014; Rommers et al., 2013) does show effects that seem to demonstrate a more unitary nature for certain types of unit, many studies that make claims about "holistic" storage of formulaic sequences in fact only show that they are processed quickly (Siyanova-Chanturia, 2015), which may be the result of a number of different underlying mechanisms (Wray, 2012). In this article, we assume that holistic or whole-form processing is a useful way of conceptualizing the widely attested processing advantage for idioms and other formulaic units, rather than necessarily indicating the discrete existence of unitary forms.

This processing advantage for idioms can be described in terms of two processes: form activation and meaning activation. Here, form activation

https://doi.org/10.1017/S0272263115000492 Published online by Cambridge University Press

refers to the recognition of specific word combinations, leading to, for example, faster initial reading of formulaic sequences, or faster responses to tasks that require a judgement of lexical form. In turn, this might be seen in facilitation for the whole phrase, or just for any portion past the "recognition point" at which an idiom is identified. Meaning activation refers to the ability to understand an intended phrasal meaning, and to integrate this into the surrounding context. This would be seen in, for example, overall reading times for sentences containing idioms, or tasks requiring a semantic judgement, such as whether a word combination is a meaningful phrase in the target language. In native speakers, then, formulaic sequences are generally privileged both in how quickly the specific word combination is recognized (e.g., Carrol & Conklin, 2014b, 2015; Siyanova-Chanturia, Conklin, & Schmitt, 2011) and in how the phrase-level meaning is processed (Swinney & Cutler, 1979; Tabossi, Fanari, & Wolf, 2009; Titone & Connine, 1999).

What Underpins the Formulaic Advantage in Idioms? There is still no clear consensus on what drives the robust idiom advantage that is apparent among native speakers. Modern theories of idiom processing have converged on a view of idioms as being simultaneously compositional *and* noncompositional/unitary. That is, a noncompositional entry for the whole unit exists at some level of representation, and this is accessible via some combination of the component words, which are assumed to be compositional/analyzable (Cacciari & Tabossi, 1988; Holsinger, 2013; Libben & Titone, 2008; Smolka, Rabanus, & Rösler, 2007; Sprenger, Levelt, & Kempen, 2006; Titone & Connine, 1999). In all of these models, subjective familiarity is seen as a key driver of faster processing—that is, an individual speaker must know a particular idiom in order to recognize it and process it quickly. Tabossi et al. (2009) showed that idioms, but also compositional phrases (clichés, such as *conquer the world*), were processed more quickly than control phrases and suggested that familiarity is the main driver of this advantage. However, despite the clear importance of subjective familiarity, idioms are often relatively infrequent, at least based on traditional corpus data (Moon, 1998), although it should be noted that as they relate to multiword sequences in general, other factors such as transitional probability and more nuanced features of word co-occurrence may complicate the picture beyond simply looking at frequency as an overall measure of phase occurrence.

Jolsvai, McCauley, and Christiansen (2013) suggested that semantic properties also contribute to fast processing in idioms. They used a phrasal decision task and asked participants to judge whether a particular three-word sequence was meaningful as an isolated unit. They compared idioms with compositional phrases and sentence fragments, with all materials matched across conditions for phrase frequency; hence the

https://doi.org/10.1017/S0272263115000492 Published online by Cambridge University Press

sequences were equally common and differed only in their meaningfulness ratings, which were assessed in a separate norming task. Frequency of occurrence facilitated processing within all three conditions, but idioms were consistently judged to be acceptable phrases more quickly than the other two conditions, suggesting that their meaningfulness contributed to faster processing. Overall, then, although formulaic language in general is processed quickly because it is frequent and familiar to native speakers, idioms demonstrate additional semantic properties that seem to contribute to their faster recognition and comprehension.

How Do Nonnative Speakers Process Formulaic Language? Three questions are important for research into how nonnative speakers process idioms in the L2: whether nonnatives show the same processing advantage as native speakers, whether L1 and L2 frequency is a key factor in processing, and whether nonnatives have a fundamentally different approach to processing in their L2.

Results of studies exploring the formulaic advantage in nonnatives are mixed, with some suggesting that the fast processing for idioms is absent in L2 speakers (Carrol & Conklin, 2014b; Siyanova-Chanturia et al., 2011). Other studies have shown, especially at higher levels of proficiency, clear effects of nonnative speakers being sensitive to L2 frequency for other formulaic sequences such as collocations (Durrant & Schmitt, 2010; Isobe, 2011; Jiang & Nekrasova, 2007; Wolter & Gyllstad, 2013). A logical assumption is that for formulaic combinations in either the L1 or L2, frequency of input or degree of exposure is a key driver of how patterns will be registered, in line with a usage-based account of language organization (Bybee, 2006; Tomasello, 2003; Wulff, 2008). This means that language-specific experience will be a strong predictor of how familiar word combinations are processed in the L1 and L2.

Importantly, for idioms not only the individual words but also an additional phrasal meaning must be learned. It has been suggested that language learners do not automatically activate the phrasal meaning (which is directly retrieved by native speakers; Titone & Connine, 1999), and instead the default position is one of compositional analysis of the literal meaning (Cieślicka, 2006, 2013; Matlock & Heredia, 2002; Siyanova-Chanturia et al., 2011). Thus on encountering a sentence like *After a long battle with cancer, my neighbor kicked the bucket*, nonnative speakers would activate the individual word meanings and might interpret the sentence as meaning that a physical bucket had been kicked, leading to difficulty comprehending the sentence as it stands. Reanalysis and consideration of the idiomatic interpretation may be possible, provided that this phrase is known in the first place. Although not a universally accepted view, various researchers have suggested that the literal meanings of individual component words are more salient to nonnative speakers and that literal interpretation of the whole phrase is therefore more

https://doi.org/10.1017/S0272263115000492 Published online by Cambridge University Press

prominent than for native speakers (Cieślicka & Heredia, 2011; Cieślicka, Heredia, & Olivares, 2014; Kecskés, 2000).

One study to examine how nonnative speakers process idiomatic word combinations is Siyanova-Chanturia et al. (2011). They compared reading times for native and highly proficient nonnative speakers on literal versus figurative uses of literally plausible idioms (e.g., *at the end of the day*). Native speakers read idioms more quickly than lower frequency control phrases like *at the end of the war*, regardless of whether the context rendered the phrase figurative or literal. Nonnative speakers, all of whom were of advanced proficiency, showed no advantage for the idioms compared to the control phrases. In addition, figurative uses showed longer overall reading times than literal uses, suggesting that the noncompositional nature of the phrases made them harder to process. Such results support the idea that nonnative speakers show a greater tendency to rely on the literal meanings of individual words in the L2, and to rely on the L1 conceptual system to try to infer a figurative meaning for a given word combination (Kecskés, 2000), or to consider idioms to be more decomposable than native speakers would (Abel, 2003). Yeganehjoo and Thai (2012) showed that this may change as proficiency develops. On a crossmodal priming task, advanced Iranian learners of English showed a greater degree of identity priming for idioms than for literal phrases (e.g., *cake* primed *The test was a piece of cake* to a greater degree than *The test was to bake a cake*). This replicates the findings of Sprenger et al. (2006) for native speakers and suggests that at high levels of proficiency and with sufficient exposure to idioms, nonnatives may start to develop nativelike representations for some phrases.

Due to less exposure and/or a more analytical approach, it seems that, in general, nonnative speakers do not show the same speeded processing of idioms in the L2 as demonstrated by native speakers, and this is true across a range of proficiency levels. In other words, known lexical combinations may not be as easily activated, and figurative meanings may not be available as early as literal meanings of words. This is not to say that idioms may not be understood, just that the mechanisms underlying their access are either qualitatively different than in the L1 or simply slower, although this may change as proficiency develops. An important related question is how well learners are able to utilize their existing L1 knowledge to aid understanding of L2 formulaic language, which is what we consider next.

Formulaic Transfer from the L1: Better the Devil You Know

In this section we consider not only idioms but also other types of formulaic language (collocations) to give a more complete picture of how

https://doi.org/10.1017/S0272263115000492 Published online by Cambridge University Press

nonnatives utilize L1 knowledge when they process words in their L2. It seems reasonable to assume that all languages contain formulaic patterns, so all language learners have a store of prefabricated word combinations in their L1 to draw on. Often idioms do cross the language barrier, likely because of the universal conceptual metaphors that underpin them in many cases, but also due to linguistic and geographical proximity and interaction. For example, German and Dutch are likely to share more idioms than either language would with Mandarin, because the languages are more closely related and because the speakers are likely to have been in closer contact throughout history.

L1 Transfer in Comprehension and Production Studies. Logically, learners should already know certain idioms in the L2 if they are congruent (have the same form and meaning in both languages). However, Kellerman (1977, 1986) demonstrated that learners are often reluctant to transfer more idiomatic senses of words, believing them to be highly language specific. In his studies of Dutch learners of English, more figurative uses of *breken* (*to break*) were rejected, even when verbatim translations of uses like *break a strike* would be acceptable in both languages. Less proficient learners showed a greater willingness to accept such transfer, whereas more advanced learners were resistant, considering idioms to be too marked and language specific to be transferable.

Contrary to this finding, subsequent studies have demonstrated that equivalence between languages can be facilitative, and often learners are very willing to transfer idioms from the L1 to aid L2 production. Irujo (1986) showed that advanced learners (Spanish L1) were able to produce significantly more English idioms via a recall and translation task when they had congruent forms in Spanish (e.g., *she wears the trousers—Ella lleva los pantalones*). Laufer (2000) found that for first- and second-year university students (L1 Hebrew) of advanced proficiency, the degree of language overlap was a clear determining factor in which idioms were correctly used in a written translation test. Total language overlap led to greater likelihood of use, but partial overlap—such as English *lip service* versus Hebrew *lip tax*—and conceptual nonequivalence, in which an idiom can only be expressed literally in the L1 (such as the English *not my cup of tea*, which has no figurative equivalent in Hebrew), were more likely to be avoided. Charteris-Black (2002) conducted a study with Malay learners (third-year English undergraduates at a Malaysian university). Students showed the greatest degree of difficulty with idioms in which there was linguistic overlap but a different conceptual meaning, or with culture-specific expressions in which no conceptual or linguistic equivalence exists in the L1. Bulut and Çelik-Yazici (2004) and Liontas (2001) showed that L2 learners utilize multiple cues and a range of strategies to understand idioms. These studies looked at advanced learners of English with L1 Turkish and L1 Greek, respectively, and found that L1

https://doi.org/10.1017/S0272263115000492 Published online by Cambridge University Press

knowledge, consideration of literal and figurative meanings, and guessing from context were all used to identify and comprehend L2 idioms. Liontas (2001) found that for both matching and nonmatching items, the addition of supporting context was facilitatory, highlighting the use of L1 knowledge, contextual clues, and more general inferencing ability in how L2 speakers are able to understand idioms in their second language.

Online Processing and L1 Transfer. Recent studies have also focused on the online processing of idioms and other types of formulaic language in the L2. For example, Titone et al. (2015) examined the effect of code-switching on sentences that contained English idioms and congruent English-French idioms. They used word-by-word presentation to show English-French and French-English bilinguals English sentences that were idioms or literal controls, and in which the final word was either English (intact condition) or French (code-switched condition). Participants then made a decision on whether each sentence was meaningful. Results suggested that code-switches during an idiom were more disruptive than during a literal sentence, but that greater congruency between languages reduced the amount of disruption. The authors proposed that this is evidence for the representation of congruent idioms in both languages, suggesting that disruption is less severe in cases of high cross-language overlap because the holistic form of the idiom exists in both English and French.

Wolter and Gyllstad (2011, 2013) employed two different methodologies to show that congruent collocations were processed more quickly than noncongruent combinations by advanced Swedish learners of English. They used a primed lexical decision task with verb + noun pairs (Wolter & Gyllstad, 2011) and a phrase-level judgement task with adjective + noun pairs (Wolter & Gyllstad, 2013). In both studies congruent items (e.g., *give an answer*, *high profile*) were judged to be acceptable more quickly and with fewer errors than incongruent (English-only) collocations (e.g., *pay a visit*, *false teeth*). Yamashita and Jiang (2010) found a similar result for Japanese-English learners, with congruent collocations judged more quickly and more accurately than incongruent ones, although this varied as a function of proficiency. They found that higher level learners showed a difference in error rates but not response times, whereas intermediate learners showed less accurate and slower responses. The authors interpreted their results as evidence that L2 exposure and L1 congruency combine to affect acquisition of formulaic patterns in nonnative speakers. Although these studies focus on collocations, they are relevant to idioms because they consider how specific word combinations are treated when encountered in a L2. The studies also include items that could be argued to be at least partially idiomatic (e.g., *broken heart*).

https://doi.org/10.1017/S0272263115000492 Published online by Cambridge University Press

Transfer and Processing of Noncongruent Forms. Other studies have specifically considered formulaic sequences in which there is a total imbalance in the L1-L2 frequency—that is, L1 formulaic items that do not exist in the L2. Carrol and Conklin (2014b) conducted a study with intermediate-proficiency Chinese learners of English to examine how translations of Chinese idioms were processed in English. They used the first few words of an idiom as a prime (e.g., *on the edge of your . . .*), and then participants made a lexical decision on idiom-completing words (*seat*) versus control words (*plate*). The study used English idioms and transliterations of figurative Chinese idioms (e.g., *draw a snake and add feet*). English native speakers showed faster responses to English idioms versus controls but showed no difference for Chinese phrases relative to controls, whereas Chinese native speakers showed no difference for English items but were consistently faster for Chinese idioms compared to controls. Similar results were found in a follow-up eye-tracking study with a similar population and rationale (Carrol & Conklin, 2015). Even for noncongruent forms, the Chinese participants in both studies showed a consistent advantage for idioms taken from their L1, despite never having seen these before in English.

Wolter and Yamashita (2014) and Ueno (2009) conducted studies looking at collocational patterns among Japanese learners (intermediate and advanced groups) and found differing results. Both studies investigated whether patterns that would be acceptable in the L1 were facilitated in the L2 (e.g., *forgive marriage*, which would be an acceptable collocation in Japanese but not in English; it is roughly equivalent to *consent to marriage* in English). Wolter and Yamashita's (2014) study used a phrase-level decision task, comparing translated L1 collocations with baseline items made up of random recombinations of experimental word pairs. They found no advantage for either adjective-noun (*bitter win*) or verb-noun (*drink tears*) combinations. Conversely, Ueno (2009) used a primed lexical decision task and did find evidence of facilitation for such combinations, but only for very advanced learners. She suggested that this was evidence that as proficiency develops, rich semantic networks are formed that encompass both the L1 and L2 in a nonselective manner. (NB: Given that other researchers—e.g., Jiang, 2000—have suggested that the role of the L1 should in general diminish as proficiency develops, this conclusion should perhaps be interpreted with caution.)

Both Ueno (2009) and Wolter and Yamashita (2014) draw on the model outlined by Jiang (2000; itself built on models first proposed by Levelt, 1989) to explain how L1 knowledge might be activated by L2 forms. In Jiang's model, all lexical entries consist of a lexeme level and a lemma level. The lexeme level, containing information about phonology, orthography, and morphology, can be roughly equated to a level of representation for form, and the lemma level, relating to semantic and

https://doi.org/10.1017/S0272263115000492 Published online by Cambridge University Press

syntactic information, to underlying meaning. Wolter and Yamashita (2014), among others, argue that the lemma-level information may also encompass aspects such as the collocational links and patterns of association that fall under the purview of formulaic language. Jiang's model suggests that the first stage of learning a language is the formation of a formal entry for a new word; hence a L2 lexeme entry is created that links to an existing L1 lemma (e.g., it is learned that the form of the French word *chien* refers to the existing L1 lexeme/lemma entry for *dog*). The second stage occurs when repeated activation of the L2 form serves to strengthen the link with the L1 lemma and to effectively copy this information into a dedicated L2 entry that remains L1-like in its makeup. A final stage involves the gradual supplanting of this L1 lemma with a more L2-like entry as a result of prolonged exposure to the L2 and represents the highest level of acquisition. However, Jiang (2000) argued that, due to the "practical constraints imposed on L2 learning" (p. 47), many words fossilize at the second stage, so even well-established L2 words may retain an underlying lemma that is fundamentally L1-like.

As it relates to formulaic language, encountering an L2 form may therefore activate lemma-level information from the L1. As Yamashita and Jiang (2010) described, encountering L1 forms should activate not only L1 translation equivalents but also L1 lexical networks. It is possible that this should therefore activate syntagmatic information about possible collocations and commonly co-occurring words (including idiom component words). Arguing against this, Williams and Cheung (2011; see also de Groot & Nas, 1991; Williams, 1994) showed that more central aspects of semantics but not associative relations showed cross-language priming. They found significant cross-language priming for translation equivalents (e.g., *squirrel*/*écureuil*) and semantically similar words (e.g., *sofa*/*chaise [chair]*), but not for semantic associates (e.g., *desk*/*chaise*).[1] They argued that associative relationships were established more through experience; hence they highlight "the importance of individual learning episodes in providing the meanings with which they are associated" (p. 93). If this view is accurate, information such as how a word combines with other words to create formulaic configurations may not form part of the core lemma-level knowledge that is linked to the L2 form but may instead be dependent on the language-specific frequency of encounter.

A final point worthy of mention is Wray's (2002) idea that components of formulaic sequences may exist multiple times in the lexicon, as discrete entries and as part of a larger "unit." Applying such a view to cross-language transfer, "core" single-word entries in the L1 lexicon may be copied to the L2, but duplicate entries that form part of larger sequences might only exist in the L1, at least until congruent forms have been encountered in the L2. Congruency between languages may

https://doi.org/10.1017/S0272263115000492 Published online by Cambridge University Press

therefore show an effect for items that do exist as duplicate entries in both languages, whereas for L1-only idioms there should be no whole form in the L2. Again, we should be careful about adopting this view of idioms and other formulaic units as whole units, given the lack of direct evidence (Siyanova-Chanturia, 2015; Wray, 2012), but as an alternative to a lemma mediation view, it is worthy of consideration.

Summary

In summary, there is clear evidence that formulaic language holds a privileged processing status for native speakers, but this is not necessarily the case for nonnatives. Native speakers process frequent, familiar word combinations quickly (a lexical/form-based advantage), and, in the case of idioms, often access the phrase-level figurative meanings as quickly or more quickly than comparable literal phrases (a meaning-based processing advantage). For nonnative speakers, L1 knowledge and L2 proficiency/exposure are both important factors in how formulaic language is processed in the L2, especially in receptive tasks in which learners can use multiple sources of information to reach a decision about the likely meaning of idioms and other phrases. It seems clear that congruency between languages can have a facilitative effect when learners encounter L2 formulaic language, but the extent of this will be determined by many factors (including the nature of the task, the perceived transferability of the item in question, and learner-specific factors like proficiency).

The present study aims to add to the literature on nonnative processing of formulaic language by exploring the importance of L1 knowledge in the online processing of idioms in the L2. The following research questions are defined:

1. Do translations of idioms show privileged processing by nonnative speakers?

This question allows us to directly test the influence of L1 patterns on how L2 word combinations are processed by nonnative speakers. Previous studies discussed in the introduction have shown mixed results; hence this will enable us to further test the extent to which L1 knowledge is used in the online processing of idioms in the L2.

2. Does congruency between languages show any additional facilitatory effects, compared to items that only exist in the L1?

This allows us to differentiate those studies that have found facilitation for congruent items (e.g., Wolter & Gyllstad, 2011, 2013; Titone et al., 2015;

https://doi.org/10.1017/S0272263115000492 Published online by Cambridge University Press

Yamashita & Jiang, 2010) and those that have looked at facilitation for translated L1-only items (e.g., Carrol & Conklin, 2014b, 2015). In other words, is cross-language facilitation purely the result of transferred L1 knowledge, or is additional experience of the same combinations in the L2 an added benefit?

3. Do advanced nonnatives show any evidence of formulaic processing for L2-only idioms?

Again, results are mixed as to whether nonnatives demonstrate frequency effects for L2 formulaic sequences. In this study we explored whether a group of high-proficiency L1 Swedes showed evidence of formulaic processing in the L2. Given the prevalence of English in Sweden and the advanced proficiency of the participants (students at an English language university in Sweden), we assumed that such a group would be most likely to demonstrate L2 formulaic processing, compared to previous studies of variable-proficiency cohorts.

METHODOLOGY

In the present study participants read idioms embedded in short, context-neutral sentences. All materials were presented in English, and we recorded the reading patterns for the whole idiom (hereafter *phrase-level measures*) and its final word (*word-level measures*). In each case we compared these to control items, created by changing the first word of each idiom to make a logical, matched alternative (e.g., *spill/drop the beans*).

English native speakers and nonnative English participants (L1 Swedish) were tested on a set of English idioms, translated Swedish idioms, and congruent idioms. We used eye-tracking to measure the number and duration of fixations during natural reading. Eye-tracking is a useful methodology for investigating the processing of formulaic units, as it enables us to consider a range of measurements and to relate these to the processes underlying the recognition and understanding of phrases in context. A central assumption in eye-tracking is that what is being looked at is a reflection of what is being processed (Pickering, Frisson, McElree, & Traxler, 2004); therefore more and longer fixations reflect greater cognitive effort. In other words, words and phrases that are easier to access/process should show shorter reading times. One challenge when applying this to formulaic language, however, is to work out how best to analyze "single" items that span several words. In this study we adopt a hybrid method of analysis (as discussed in Carrol & Conklin, 2014a), whereby we consider a range of early and late measures at both the word level and the phrase level (see Table 1). Early measures are

https://doi.org/10.1017/S0272263115000492 Published online by Cambridge University Press

Table 1. Eye-tracking measures, along with descriptions and stage of processing

Stage of processing	Type of measure	Description
Phrase Level		
Early	First pass reading time	The sum duration of all fixations on the phrase the first time it is encountered in the sentence
Late	Total reading time	The sum duration of all fixations on the phrase during the trial (including rereading)
	Total fixation count	The total number of fixations on the phrase during the trial
Word Level		
Early	Likelihood of skipping	The likelihood that a word is skipped (not fixated on at all) during first pass reading
	First fixation duration	The duration of the first fixation on the word
	Gaze duration	The sum duration of all fixations on the word the first time it is encountered in the sentence
Late	Total reading time	The sum duration of all fixations on the word during the trial (including rereading)
	Regression path duration[a]	The sum of all durations from first fixation on a word until leaving to the right, including all regressive fixations on preceding words

[a] Regression path duration is sometimes seen as an early measure and sometimes as a late measure (Clifton, Staub, & Rayner, 2007). We consider it to be a late measure here because it is likely to reflect difficulty integrating the final word into the overall phrase, leading to reconsideration of the preceding context in order to resolve the difficulty.

generally taken to reflect lexical access and other automatic processes, whereas late measures are seen as reflecting postlexical strategic effects (Altarriba, Kroll, Sholl, & Rayner, 1996; Inhoff, 1984; Paterson, Liversedge, & Underwood, 1999; Staub & Rayner, 2007) but may also be indicative of other processes—for example, if there is conflict with the preceding context. In the current study, we can relate this to the distinction between form and meaning activation: Early measures can be seen to reflect how easily the expected lexical combinations are activated, whereas later measures show how easily the overall meaning is activated and integrated into the wider sentence (including whether this requires any reassessment of the prior context).

https://doi.org/10.1017/S0272263115000492 Published online by Cambridge University Press

Participants

Twenty-four English native speakers and 24 Swedish native speakers took part in the study and received a fee for their participation. Native English speakers were all undergraduates at a U.K. university with L1 English and no experience of learning Swedish. Nonnative English speakers were all students at an English language university in Sweden. Most were undergraduates (there was one postgraduate) and were studying English language and literature. All had Swedish as their L1. The entry requirements for these learners in terms of English proficiency correspond to either an IELTS score of at least 6.5 (academic), a TOEFL result of at least 575 points (paper-and-pencil test) or 90 points (Internet-based test), or a Certificate in Advanced English (CAE) from Cambridge English Language Assessment. Following the main experiment, demographic and language background data were collected, including self-rating of proficiency in English and an estimate of usage in various contexts (e.g., at the university, at home with friends and family, reading for pleasure, etc.). A short vocabulary test was also administered, consisting of a shortened version of the Vocabulary Size Test (Nation & Beglar, 2007).[2] In this test items are presented in a short, neutral context (e.g., Shoe: This is a shoe), and participants select the correct definition from four alternatives; we added a "Don't know" option to minimize guessing, as per the suggestion in Zhang (2013). The original test sampled 10 items from each of the first 14 British National Corpus (BNC) (Davies, 2004) word levels (Level 1 represents the 1,000 most frequent word families in English, Level 2 the next 1,000, and so on). We randomly selected 2 items from the first 10 bands to give a total of 20 items; thus a score of 20/20 would correspond to a vocabulary size of around 10,000 words. The mean score on this test was 16.2/20 (*SD* = 2.4, range = 11–20, reliability [Cronbach's alpha] = .77[3]). This corresponds to around 8,000 word families, which was in keeping with previous studies of typical vocabulary sizes among Swedish undergraduate university students (Gyllstad, 2007, 2012). We also assume that vocabulary size is a reliable proxy for language proficiency overall (Alderson, 2005; Meara & Jones, 1988). As reported in Wolter and Gyllstad (2013), there is no universally agreed on measure of what constitutes intermediate or advanced proficiency, but these authors cite Milton (2010), who suggests that attaining the highest levels of C1/C2 on the Common European Framework for Languages is associated with approximate receptive vocabulary scores of 3,750–5,000 words. Other estimates vary; for example, Nation (2006) suggests that 8,000–9,000 word families are required to understand written texts (newspapers and novels), and that 6,000–7,000 are required for spoken comprehension. All learners in the present study exceeded the threshold of 5,000 word families, and the majority (21/24) showed

https://doi.org/10.1017/S0272263115000492 Published online by Cambridge University Press

scores that can be extrapolated to reflect a vocabulary size of at least 8,000 word families. We therefore consider the nonnative participants in this study to be a fairly homogeneous cohort of advanced learners of English. A summary of the nonnative participants is provided in Table 2.

Materials

Three categories of stimuli were created: English-only idioms, Swedish-only idioms, and congruent idioms—idioms with the same or very similar form and meaning in both languages. All were selected to conform as closely as possible to the structure X-det-N, in which X was a verb (e.g., *kick the bucket*) or in some cases a noun (*neck over head*) or preposition (*under the ice*). The determiner was sometimes a personal pronoun (e.g., *pull your weight*), was sometimes replaced by a preposition (*fall from grace*), or was sometimes omitted (*tread water*). The key criterion was that each item must contain two main lexical items, and some flexibility was permitted to ensure that sufficient numbers of items could be found in each of the three categories. Many previous studies have used idioms of variable length (e.g., Carrol & Conklin, 2014b; Titone et al., 2015); however, in these cases predictability can be a potentially confounding factor, meaning that English native speakers will be likely to actively guess the completion to phrases like *flog a dead . . .* (*horse*). By using only very short idioms, we aimed to minimize this kind of guessing. All experimental items are available in Appendix A.

English idioms were first selected from a variety of sources, including from previous studies by the authors and various idiom dictionaries (principally Warren, 1994). An initial pool of around 100 common English idioms was prepared. This list was examined by one of the authors, a Swedish native speaker, who identified all idioms that have a corresponding version with identical or near-identical form in Swedish—for

Table 2. Summary of nonnative speakers (L1 Swedish), n = 24

	Age	Years of English	Reading	Listening	Speaking	Writing	Usage	Vocab.
Mean	23.7	11.5	7.4	8.1	7.0	7.0	39.5	16.2
SD	5.9	2.3	1.5	1.2	1.3	1.6	5.8	2.4
Range	19–45	9–19	4–10	5–10	4–9	4–10	29–49	11–20

Note. Years of English is the amount of formal instruction each participant had undergone at the time of testing; reading, listening, speaking, and writing are self-ratings out of 10; usage is an aggregated estimate of how often participants use English in their everyday lives (10 measures, each estimated out of 5, to give a total score out of 50); and vocab is a score out of 20 on the modified 1–14K English vocabulary size test.

https://doi.org/10.1017/S0272263115000492 Published online by Cambridge University Press

example, *break the ice*, which has a direct equivalent *bryta isen*. This judgement was based on personal experience and was checked using a variety of Swedish idiom dictionaries and lists (principally Hargevik & Ljung, 1989; Hübinette & Odenstedt, 1988). In all cases the main lexical items had single-word translation equivalents and appeared in the same order in both languages, although because Swedish definite articles are attached to the end of the noun they modify, some variation in form was inevitable (e.g., *ice* = *is*, *the ice* = *isen*). Final sets of 40 idioms were created for each condition (congruent and English-only), with certain items discounted if they included very low frequency vocabulary items.

A final list of Swedish-only idioms was prepared by the Swedish author. These consisted of idioms of the same general form: two main lexical items, mostly V-det-N but also, in some cases, N-Prep-N (*a cow on the ice*) and det-Adj-N (*the red thread*). The majority of idioms in this condition (around 80%) conformed to the V-det-N structure. All were chosen from various Swedish idiom dictionaries and word lists, as before. The list was reviewed by the native English-speaking authors to ensure that none of the idioms existed in English. These were then transliterated into English as closely as possible, with the core meaning of each word taken as the basis for translation by the Swedish author. These translations were checked using Google Translate and then submitted to a translation norming test using Swedish native speakers who were advanced learners of English (either lecturers in English or, in one case, a postdoctoral researcher; thus their proficiency was nativelike or near-nativelike). They were asked to assess the English translations for accuracy using a 5-point scale and, where appropriate, suggest any improvements. Overall ratings were high (mean = 4.7/5, *SD* = 0.4, range = 3.0–5.0), and any items that received scores below 4/5 were amended as per the suggestions given by the raters. These suggestions were generally very minor (e.g., *neck* instead of *throat* for the item *hals över huvud [neck over head]*).

All idioms were presented in a short norming study to assess how well known they were to native speakers of English. Participants (n = 13) were asked to indicate familiarity with each phrase on a 7-point scale (1 = completely unfamiliar, 7 = very familiar). English-only items and congruent items were very well known, whereas translations of Swedish items were unknown (see Table 3; NB: Ratings were collected for the idioms only, not the literal control items). Subjective familiarity ratings for the idioms were also collected following the main experiment on a by-subject basis—that is, participants were asked to rate their own personal familiarity with each experimental item.

For all idioms a control phrase was created by changing the first content word to an alternative matched for part of speech and, where possible, length and frequency—for example, *break the ice* became *crack the ice*. All control phrases therefore formed logical, acceptable, but nonidiomatic

https://doi.org/10.1017/S0272263115000492 Published online by Cambridge University Press

Table 3. Summary of item characteristics for all idioms and control phrases

	Swedish		Congruent		English	
	Idioms	Controls	Idioms	Controls	Idioms	Controls
Phrase length (characters)	14.2 (3.6) 8–22	13.8 (3.4) 7–20	14.5 (1.9) 10–18	14.6 (2.0) 10–19	14.4 (2.2) 10–19	14.3 (2.3) 10–20
Phrase freq.[a]	23 (132) 0–834	34 (188) 0–1,207	56 (42) 6–193	6 (8) 0–35	40 (38) 10–224	7 (12) 0–71
Familiarity	1.6 (1.0) 1–5.4	n/a	6.0 (1.0) 2.3–6.9	n/a	6.2 (0.8) 3.2–6.9	n/a
Word 1 length (letters)	5.2 (1.7) 3–10	4.7 (1.3) 2–7	4.9 (1.5) 3–9	4.9 (1.6) 3–9	4.7 (1.2) 3–7	4.8 (1.3) 3–8
Word 1 freq.	46,227 (63,935) 294–211,009	59,008 (105,839) 379–643,901	59,897 (205,448) 121–1,304,998	39,403 (46,867) 653–176,925	31,969 (48,060) 961–208,322	73,921 (207,764) 1,271–1,304,940
Word 2 length (letters)	5.2 (1.6) 3–8		4.9 (1.0) 3–7		4.8 (1.2) 3–8	
Word 2 freq.	7,425 (11,072) 120–47,353		23,672 (40,324) 68–175,076		14,692 (17,476) 791–90,846	

Note. Table displays mean values (*SD*s in brackets), with range underneath. Phrase length includes spaces; phrase frequency and Word 1 or 2 frequency (BNC) are per 100 million words; and familiarity is an average rating on a 7-point scale based on prenorming.

[a] Phrase frequency for Swedish items was hugely inflated by the high occurrence of *gå bort* (*walk away*, meaning "to die"), which occurs in its literal form in English 834 times, as well as in its control form, *move away* (frequency = 1,207). Without this item, mean phrase frequency for Swedish idioms was 3 and for controls was 4. This item was retained on the grounds that it is not an idiom in English, despite its high frequency.

sequences in English. Short sentence contexts were then created for each item. Context can be an important factor in the processing of different kinds of idioms (e.g., Cieślicka, 2013; Titone & Connine, 1999), with a biasing context greatly increasing predictability. We therefore ensured that all contexts were created to be neutral, that is, that they did not bias a figurative or literal meaning of the idiom (see Appendix B for examples). Thus, encountering the first word (e.g., *kick* in *kick the bucket*) would not lead participants to expect an idiom completion any more than they might expect a literal completion. The context was created so that all literal control phrases were logical and grammatical, but the idioms varied according to whether the figurative meaning was known. For translated Swedish items, this meant that the contexts would only be grammatical if the idiom was understood in its figurative sense. Hence a phrase like *hot on the porridge*, meaning "over eager," is only grammatical/logical in English if the figurative meaning is known, in the same way that otherwise ungrammatical phrases in English are acceptable when used in certain contexts, such as *by and large* or *long time no see*. Similarly, English idioms would only be considered logical by Swedish native speakers if the figurative meaning were known (as in the example in [1] below, in which not knowing the figurative meaning would render the sentence semantically anomalous).

In all cases the material preceding and immediately following the idiom/control phrase was the same for both versions, for example:

(1) Idiom sentence: It was hard for him to *break the ice* when he was at the party last week.
(2) Control sentence: It was hard for him to *crack the ice* when his locks froze last week.

Idioms/controls were therefore matched for number of preceding words (for both idioms and controls, mean = 4.0, *SD* = 0.8, and range = 2–6) and were comparable as to the number of words following the phrase (idioms: mean = 11.2, *SD* = 1.8, range = 8–17; controls: mean = 11.8, *SD* = 1.9, range = 7–16). By creating control phrases in which the first word of each phrase—rather than the terminal word—was changed, we could directly compare reading times for the same word in different contexts, rather than comparing different words, as has often been the case in previous idiom studies (e.g., *break the ice* vs. *break the cup*).

Table 3 provides a summary of the distributional properties (length and frequency) of the idioms and control items, for both phrases and component words. Note that because the control items were created by changing the first word of each phrase, values for the final word are identical between idioms and controls in each condition.

Counterbalanced lists were created to ensure that no participant would see both the idiom and the corresponding control phrase in the

https://doi.org/10.1017/S0272263115000492 Published online by Cambridge University Press

same study. Lists were matched internally (across conditions) and externally (relative to each other) for phrase frequency, and for length and frequency of the individual words. A number of filler sentences were included so that, overall, only 25% of sentences contained an idiom.

Procedure

The study was conducted using an Eyelink 1000+ system for the English native speakers and an Eyelink 1000 system for the Swedish speakers. Recording was performed with a desk-mounted eye-tracking camera and was monocular at a sampling rate of 500 Hz. Participants were seated in front of a 1,280 × 1,960 resolution widescreen monitor with a refresh rate of 144 Hz. Head position was stabilized with a desk-mounted chin rest.

Participants were randomly assigned to one of the stimulus lists. An initial instruction screen was shown, followed by camera setup and calibration. Participants were shown five practice trials, followed by the experimental items. At the start of each trial a fixation cross appeared toward the center-top of the screen and then each sentence appeared on one line across the middle of the screen in Courier New font, size 18 pt. Participants were asked to read each sentence as naturally as possible for comprehension and to press the space bar as soon as they had finished reading. One third of the items were followed by a simple yes/no question, which was included to ensure that participants were actually reading for comprehension rather than just skimming the sentences.[4] The remainder of the sentences were followed by a "Ready?" prompt. Participants saw the stimulus items in two blocks of 60 sentences, with a short break after Block 1. Each block was balanced across conditions, and within each block the trial order was randomized for each participant. Trial-by-trial drift correction was monitored throughout and recalibration was performed as required. The eye-tracking took around 30 min for Swedish participants and around 20 min for English native speakers.

All participants were asked to complete a rating questionnaire to indicate subjective familiarity with the idioms used (administered after the main experiment). They were asked to judge each idiom (whether they had seen it before and whether they knew the figurative meaning) on a 7-point Likert scale. For English native speakers, all 120 idioms were presented in English in a random order. For nonnative English speakers, two versions were used. One presented the English-only and the congruent idioms in English, and the second presented the Swedish-only and the congruent idioms in Swedish. In both cases the order of presentation was randomized, and to minimize repetition effects for the congruent idioms (which appeared on both lists but in different languages), half of

https://doi.org/10.1017/S0272263115000492 Published online by Cambridge University Press

the participants saw the English list first and half saw the Swedish list first. Participants were specifically asked to indicate their familiarity with the idioms in the language in which they appeared. Finally, all participants were asked to provide some background information. For English native speakers this consisted of basic information such as age and study status. For nonnative English speakers this included a more detailed background questionnaire and vocabulary test, as described earlier and as summarized in Table 2.

RESULTS

Prior to analysis, all eye-tracking data were checked for missing or unusable trials. Any trials in which track loss occurred were removed, although this accounted for a very small fraction of all data (less than 0.01%). Data were cleaned according to the four-stage process within the Eyelink Data Viewer software. All fixations shorter than 100 ms or longer than 800 ms were removed. Fixation data were extracted for all trials for the whole phrase and for the final word of each idiom/control. Results were analyzed using an omnibus linear mixed-effects model using the lme4 package (Version 1.0–7; Bates et al., 2014) in R (Version 3.1.2; R Development Core Team, 2014). Three treatment-coded main effects of Group (English L1 vs. Swedish L1), Phrase Type (literal phrase vs. idiom), and Condition (congruent vs. English vs. Swedish) were included, as were random intercepts for subject and item and by-subject random slopes for the effects of Phrase Type and Condition, following the advice of Barr, Levy, Scheepers, and Tily (2013) to include a maximal random effects structure wherever this is justified by the experimental design. In all models we included the covariates of word length and log-transformed word frequency (for Word 1 and Word 2 for phrase-level models and Word 2 only for word-level models) to ensure that effects of these were controlled. A summary of the raw results is shown in Table 4.

Table 5 (phrase level) and Table 6 (word level) show the omnibus mixed-effects analysis for all eye-tracking measurements. All duration measures are log-transformed to reduce skewing. In all cases we report the model structure and the coefficient (β), standard error (*SE*), and *t* value (*z* value for likelihood of skipping and fixation count), along with estimated significance values. For word-level analysis, the likelihood of skipping was analyzed with a logistic mixed-effects model, and skipped items were removed from the analysis for subsequent durational measures. (See the Supplementary Materials for more information on how to interpret these models.) For simplicity, we describe and explain the important features of our results in terms of the effect of Phrase Type (do idioms show shorter reading times than controls?) and Condition (do congruent, English-only, and Swedish-only idioms

https://doi.org/10.1017/S0272263115000492 Published online by Cambridge University Press

Table 4. Results for all speakers, split by participant group and by phrase-/word-level measures

	Swedish-only		Congruent		English-only	
	Idioms	Controls	Idioms	Controls	Idioms	Controls
Swedish Native Speakers: Whole Phrase						
First pass reading time	625	670	597	596	564	609
Total reading time	1,176	1,309	997	1,062	977	1,021
Fixation count	5.0	5.4	4.2	4.6	4.2	4.4
Swedish Native Speakers: Final Word						
Likelihood of skipping	.08	.02	.13	.04	.13	.13
First fixation duration	237	256	211	229	215	207
Gaze duration	282	299	237	250	235	247
Total reading time	455	535	349	378	329	348
Regression path duration	739	867	524	617	507	531
English Native Speakers: Whole Phrase						
First pass reading time	450	463	361	415	367	430
Total reading time	832	652	475	561	466	582
Fixation count	3.9	3.0	2.4	2.7	2.3	2.8
English Native Speakers: Final Word						
Likelihood of skipping	.10	.11	.29	.25	.33	.23
First fixation duration	202	197	149	161	135	166
Gaze duration	223	208	150	166	140	173
Total reading time	337	248	179	213	159	216
Regression path duration	541	360	211	278	199	291

Note. For duration measures, reading times in milliseconds are reported; fixation count is a raw value; and likelihood of skipping is reported as a probability.

show different patterns?), for each of the participant groups (English native speakers and Swedish native speakers). Interactions among these variables would indicate differential processing according to the origin of the phrase; for example, an interaction of Group, Phrase Type, and Condition (English vs. Swedish) would suggest that English native speakers process English idioms but not Swedish translations more quickly than controls, whereas Swedish native speakers show the complementary pattern (faster processing for Swedish idioms compared to controls, but no effect for English idioms).

The omnibus analysis shows clear effects of Group, whereby English native speakers were faster readers than Swedish native speakers. There was also an overall effect of Type for most measures, which shows that, in general, participants read the idioms more quickly, fixated on them fewer times, and skipped the final words more often than they did for the literal control phrases. For all measures except likelihood of skipping the final word and first fixation duration on the final word, this effect was qualified by an interaction between Group, Type, and Condition: Swedish.

https://doi.org/10.1017/S0272263115000492 Published online by Cambridge University Press

Table 5. Omnibus mixed-effects model estimates for all phrase-level measures

	First pass reading time			Total reading time			Fixation count		
	β	*SE*	*t*	β	*SE*	*t*	β	*SE*	*z*
Fixed Effects									
Intercept	6.10	0.18	33.77	6.40	0.20	31.57	1.13	0.20	5.54
Group: Swedish	0.29	0.07	3.96***	6.46	0.09	7.56***	0.52	0.08	6.29***
Type: Idiom	–0.12	0.05	–2.56*	–0.13	0.04	–3.20**	–0.12	0.06	–2.06*
Condition: English	0.06	0.06	0.96	0.04	0.06	0.76	0.03	0.07	0.40
Condition: Swedish	0.04	0.06	0.74	0.13	0.06	2.10*	0.09	0.07	1.37
Group*Type	0.16	0.07	2.37*	0.04	0.06	0.68	0.03	0.07	0.40
Group*Condition: English	0.02	0.07	0.27	–0.08	0.06	–1.44	–0.09	0.07	–1.26
Group*Condition: Swedish	0.01	0.07	0.14	0.03	0.06	0.52	0.02	0.07	0.28
Type*Condition: English	–0.04	0.07	–0.60	–0.07	0.05	–1.24	–0.06	0.08	–0.80
Type*Condition: Swedish	0.10	0.07	1.54	0.35	0.05	6.32***	0.36	0.08	4.79***
Group*Type*Condition: English	–0.08	0.09	–0.85	0.11	0.08	1.48	0.11	0.10	1.07
Group*Type*Condition: Swedish	–0.18	0.01	–1.91+	–0.35	0.08	4.58***	–0.35	0.10	–3.64***
Control Predictors									
Word 1 length	0.01	0.01	0.97	–0.01	0.01	–0.86	–0.01	0.01	–1.03
Word 1 frequency (log)	–0.02	0.01	–2.05*	–0.01	0.01	–1.03	–0.01	0.01	–1.11
Word 2 length	0.03	0.01	2.03*	0.02	0.02	1.14	0.02	0.02	1.36
Word 2 frequency (log)	–0.02	0.01	–1.85+	–0.02	0.01	1.48	–0.01	0.01	–1.04

Continued

Table 5. Continued

	First pass reading time			Total reading time			Fixation count		
	β	*SE*	*t*	β	*SE*	*t*	β	*SE*	*z*
Random effects	Variance			Variance			Variance		
Item	0.021			0.038			0.030		
Subject	0.037			0.070			0.053		
Subject | Type	0.003			0.004			0.001		
Subject | Condition: English	0.002			0.001			0.002		
Subject | Condition: Swedish	0.006			0.004			0.004		
Residual	0.258			0.178			n/a		

Note. For condition, Congruent is taken as the baseline. Table displays coefficients (β), standard errors (*SE*s), and *t* values (*z* values for fixation count, for which a generalized linear model with Poisson regression was used). Significance values are estimated by the lmerTest package in R (Version 2.0–11; Kuznetsova, Brockhoff, & Christensen, 2014).

*** $p < .001$.

** $p < .01$.

* $p < .05$.

$^{+}$ $p < .10$.

Table 6. Omnibus mixed-effects model estimates for all word-level eye-tracking measures

	Likelihood of skipping			First fixation duration			Gaze duration			Total reading time			Regression path duration		
	β	*SE*	*z*	β	*SE*	*t*	β	*SE*	*t*	β	*SE*	*t*	β	*SE*	*t*
Fixed Effects															
Intercept	–0.43	0.73	–0.59	5.40	0.08	70.48	5.43	0.10	55.35	5.49	0.14	38.53	5.72	0.19	30.28
Group: Swedish	–2.30	0.41	–5.56***	0.10	0.05	2.16*	0.15	0.05	2.90**	0.32	0.08	4.22***	0.51	0.12	4.39***
Type: Idiom	0.22	0.22	1.02	–0.05	0.03	–1.35	–0.07	0.04	–1.73⁺	–0.11	0.05	–2.11*	–0.18	0.07	–2.81**
Condition: English	–0.19	0.26	–0.72	–0.01	0.04	–0.33	–0.01	0.04	–0.16	–0.03	0.06	–0.47	–0.01	0.07	–0.10
Condition: Swedish	–0.99	0.30	–0.72	0.03	0.04	0.85	0.03	0.04	0.73	0.02	0.06	0.27	0.11	0.07	1.54
Group*Type	1.18	0.44	2.66**	0.06	0.05	1.26	0.09	0.05	1.71	0.06	0.07	0.93	0.06	0.09	0.68
Group*Condition: English	1.57	0.45	3.51***	0.03	0.05	0.61	0.06	0.05	1.26	0.03	0.07	0.48	–0.05	0.09	–0.61
Group*Condition: Swedish	0.18	0.65	0.28	0.04	0.05	0.98	0.08	0.05	1.59	0.24	0.07	3.59***	0.18	0.09	2.04*
Type*Condition: English	0.40	0.31	1.30	–0.01	0.05	–0.21	–0.00	0.05	–0.01	–0.01	0.07	–0.17	–0.00	0.09	–0.02
Type*Condition: Swedish	–0.32	0.37	–0.86	0.06	0.05	1.23	0.11	0.05	2.16*	0.34	0.07	4.95***	0.51	0.08	6.02***
Group*Type*Condition: English	–1.86	0.57	–3.27**	0.01	0.06	0.17	–0.07	0.07	–0.96	0.00	0.10	0.02	0.08	0.12	0.65
Group*Type*Condition: Swedish	0.56	0.77	0.73	–0.08	0.06	–1.25	–0.14	0.07	–2.04*	–0.40	0.09	–4.27***	–0.47	0.12	–4.01***

Continued

Table 6. Continued

	Likelihood of skipping			First fixation duration			Gaze duration			Total reading time			Regression path duration		
	β	*SE*	*z*	β	*SE*	*t*	β	*SE*	*t*	β	*SE*	*t*	β	*SE*	*t*
Control Predictors															
Word 2 length	–0.24	0.07	–3.40***	0.00	0.00	0.18	0.02	0.01	2.61*	0.04	0.01	3.01**	0.02	0.02	1.41
Word 2 frequency (log)	0.03	0.06	0.56	–0.01	0.01	–1.92+	–0.02	0.01	–2.96**	–0.02	0.01	–1.89*	–0.02	0.01	–1.39
Random effects	Variance			Variance			Variance			Variance			Variance		
Item	0.343			0.003			0.007			0.017			0.029		
Subject	0.390			0.013			0.015			0.043			0.120		
Subject | Type	n/a			0.000			n/a			0.002			0.004		
Subject | Condition: English	n/a			0.002			n/a			0.004			0.009		
Subject | Condition: Swedish	n/a			0.002			n/a			0.005			0.007		
Residual	n/a			0.099			0.124			0.220			0.347		

Note. For condition, Congruent is taken to be the baseline. Table displays coefficients (β), standard errors (*SE*s) and *t* values (*z* values for likelihood of skipping), with significance values estimated by the lmerTest package in R (Version 2.0–11; Kuznetsova et al., 2014).
*** $p < .001$.
** $p < .01$.
* $p < .05$.
+ $p < .10$.

This suggests that whereas Swedish native speakers treated both congruent and Swedish idioms as known, English native speakers showed a significant difference in how they read these two types. Overall, the interactions suggest that the two speaker groups did show different patterns for the different conditions (English, Swedish, and congruent idioms), so to further explore the data, separate models were fitted for the Swedish native speaker and the English native speaker data (for both groups the L1-only condition was taken as the baseline—i.e., for Swedish native speakers, Swedish idioms were the baseline, so the effect of Condition: Congruent and Condition: English were considered). Interactions were explored using the Phia package (Version 0.1–5; De Rosario-Martinez, 2013) in R to conduct pairwise comparisons as appropriate. Significant results are described here, and full model outputs are provided in the Supplementary Materials (Tables A, B, and C).

Swedish Native Speakers

Swedish native speakers showed a pattern of overall facilitation for idioms compared to controls in all three conditions. At the phrase level, they showed no effects for first pass reading time but spent significantly less time overall reading idioms than controls (t = –2.65, p = .009) and also showed fewer fixations (z = –1.96, p = .051). For word-level analysis, likelihood of skipping was significantly higher for idioms overall (z = 2.96, p = .003), and there was an interaction of Type and Condition: English (z = –2.74, p = .006). This suggests that Swedish and congruent idioms showed an advantage compared to literal controls, whereas English idioms/controls showed no difference. Pairwise comparisons confirmed that the final words of idioms were skipped more often than those of the controls in the Swedish-only condition ($\chi^2(1, 1{,}434) = 8.78$, p = .006) and congruent condition ($\chi^2(1, 1{,}434) = 12.49$, p = .001) but not the English-only condition ($\chi^2(1, 1{,}434) = 0.04$, p = .84). Other early measures (first fixation duration and gaze duration) showed no significant effects. Total reading time showed an overall effect, such that idioms in all conditions were read more quickly than controls (t = –2.27, p = .024). Regression path duration showed no effects of Phrase Type, so there was no difference in encountering either an idiom or a control phrase (from any condition) in terms of having to return to the prior context to reassess and reintegrate the phrase. Importantly, there was no evidence that congruent idioms were processed any differently than Swedish-only ones. No interactions were observed between Phrase Type and Condition: Congruent for any of the phrase- or word-level measures, indicating that Swedish native speakers processed Swedish and congruent items in a similar manner.

https://doi.org/10.1017/S0272263115000492 Published online by Cambridge University Press

English Native Speakers

English native speakers showed a clear pattern across all measures except for word-level first fixation duration and gaze duration (although it should be remembered that these are strongly affected by the removal of any skipped items). As expected, there was no interaction between Phrase Type and Condition: Congruent, demonstrating that to English native speakers there was no difference between these conditions and that all items were treated as known phrases. Pairwise comparisons confirmed that for all phrase-level measures and late word-level measures (total reading time and regression path duration), they spent less time on English and congruent items compared to matched literal phrases (all *p*s < .05). Swedish idioms showed disruption on a range of measures, as indicated by the interaction of Phrase Type and Condition: Swedish for phrase-level first pass reading ($t = 2.57$, $p = .010$), total reading time ($t = 7.22$, $p = .000$), and fixation count ($z = 5.56$, $p = .000$), and on all word-level measures except first fixation duration: likelihood of skipping ($z = -1.91$, $p = .05$), dwell time ($t = 2.36$, $p = .018$), total reading time ($t = 5.85$, $p = .000$), and regression path duration ($t = 6.74$, $p = .000$). This suggests that English native speakers had difficulty with the Swedish idioms when they were first encountered, and in making sense of them in the context of the whole sentence/integrating the overall phrasal meaning. For English-only and congruent items, even though the literal control items were all perfectly plausible, there was a consistent advantage for idioms on all measures, as predicted by the previous literature. Figures 1 and 2 demonstrate the different patterns for English native speakers and Swedish native speakers on the likelihood of skipping the final word (word-level early measure) and phrase-level reading time (phrase-level late measure).

Familiarity

A set of models was fitted to assess the effect of subjective familiarity. We analyzed this separately, as different rating sets were used for the English and Swedish native speakers (detailed subsequently). Separate models were created for English native speakers and Swedish native speakers, with the interaction between familiarity rating and type (idiom vs. literal phrase) computed for each measure.

For English native speakers, the English (mean = 6.2/7, *SD* = 0.90) and congruent (mean = 6.0/7, *SD* = 1.13) categories were collapsed into one, and Swedish idioms were discounted on the grounds that they were all fundamentally unknown (mean = 1.6/7, *SD* = 1.02). English native speakers showed significant interactions between familiarity and Phrase Type for

https://doi.org/10.1017/S0272263115000492 Published online by Cambridge University Press

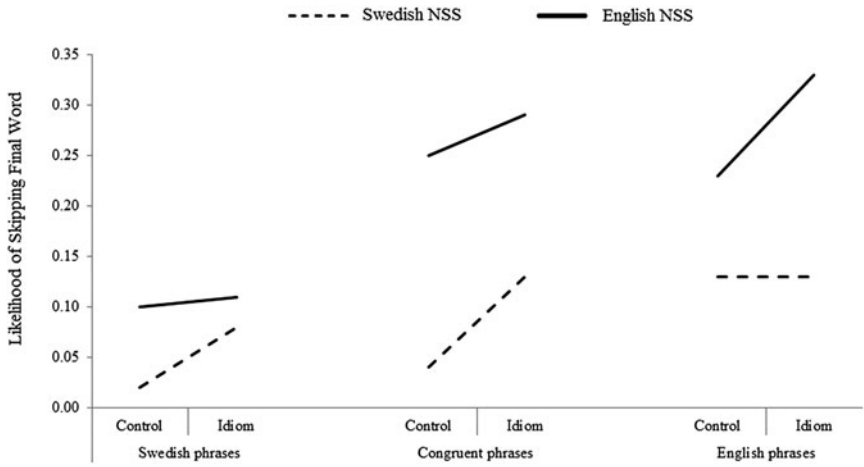

Figure 1. Interaction plots for likelihood of skipping the final word. Upward slopes indicate greater likelihood of skipping in idioms compared to control phrases. Solid lines show that native speakers were significantly more likely to skip the final word in English and congruent idioms, but showed no difference for Swedish idioms. Dotted lines show that non-native speakers showed the same pattern of skipping for Swedish and congruent items (idioms skipped more often than control phrases), but no difference for English-only idioms.

phrase-level total reading time (t = –3.32, p < .001) and word-level regression path duration (t = –2.53, p = .012); in both cases greater familiarity led to shorter reading and rereading times for English idioms. No early measures showed any effect of familiarity.

For Swedish native speakers, the effects on each condition were considered separately; for congruent items both Swedish ratings (mean = 5.4/7, SD = 0.97) and English ratings (mean = 5.7, SD = 0.94) of familiarity were considered. Swedish-only items (mean = 5.1/7, SD = 1.32) showed no effects for early measures, but there was a significant interaction between Phrase Type and familiarity for phrase-level total reading time (t = –1.97, p = .049), a marginal interaction with word-level total reading time (t = –1.74, p = .08), and a significant interaction with regression path duration (t = –2.10, p = .036). Familiarity with the L1 idiom, therefore, leads to less time being spent on the English translation for late measures, suggesting that the meaning could be more easily understood the better the idiom was known in the L1. For congruent items there were no effects of English familiarity ratings on any measure; however, for the Swedish familiarity ratings there were marginal or significant interactions with Phrase Type for phrase-level total reading time (t = –1.86, p = .060), word-level total reading time (t = –1.99, p = .047), and regression path duration (t = –1.89, p = .059). Congruent items were

https://doi.org/10.1017/S0272263115000492 Published online by Cambridge University Press

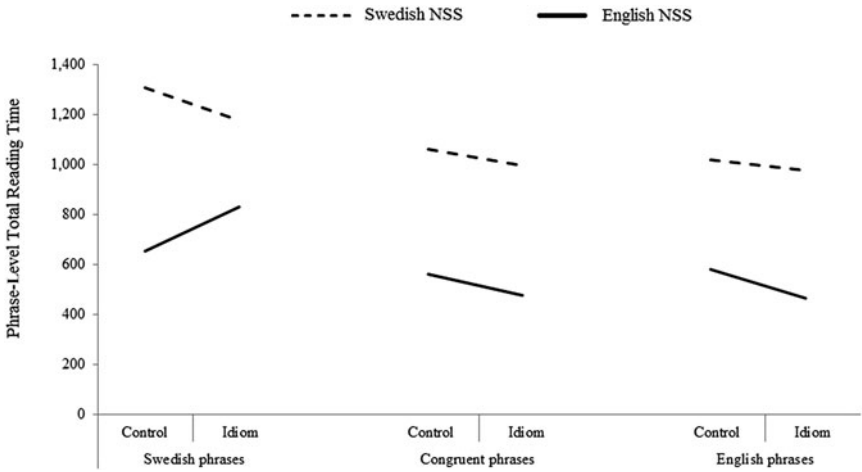

Figure 2. Interaction plots for phrase level reading times. Downward slopes indicate shorter overall reading times for idioms compared to control phrases. Dotted lines show that for non-native speakers, idioms in all three categories showed shorter overall reading times than control phrases. Solid lines show that for native speakers, English and congruent idioms showed shorter reading times, but Swedish idioms were read for substantially longer overall.

therefore affected positively by L1 familiarity for late measures (increased familiarity was facilitatory), just like Swedish-only items, but showed no evidence that specific L2 familiarity was important. No effect on early measures for either set of ratings was demonstrated. For English-only items (mean = 4.9/7, *SD* = 1.19) there were no effects of familiarity on early duration measures (phrase-level first pass reading time, word-level first fixation duration, and gaze duration); however, there were significant interactions between Phrase Type and familiarity for phrase-level total reading time ($t = -3.58$, $p < .001$), likelihood of skipping the final word ($z = 2.57$, $p = .010$), word-level total reading time ($t = -3.23$, $p = .001$), and regression path duration ($t = -3.98$, $p < .001$). For items that only exist in the L2, specific L2 familiarity is a strong predictor of how easily the idiom will be understood, and also of whether the final word is predictable enough to be skipped (whether the form of the idiom is known).

Overall, familiarity showed consistent effects in late measures only. For English native speakers this suggests that better known idioms were more easily understood, but this was not reflected in the automatic activation of known lexical combinations (no effect for early measures). This may be a reflection of the overall familiarity of the items, which were deliberately selected to be generally well known. For nonnative English speakers, L1 familiarity seemed to play some role in how both

https://doi.org/10.1017/S0272263115000492 Published online by Cambridge University Press

congruent and translated Swedish idioms were processed. A clearer finding is that when no L1 knowledge was available, in the case of English-only idioms, specific L2 familiarity was a strong factor in how stimuli were processed, consistent with the use of multiple strategies by learners discussed in the Introduction.

Proficiency

Finally, we considered the effect of individual participant proficiency by constructing models to take into account length of time learning English, aggregated self-rating scores, usage scores, and vocabulary test scores. Although higher proficiency measures were indicative of faster reading times in general for both phrase-level and word-level reading, there were no interactions with Phrase Type (idioms vs. control) or Condition (Swedish vs. English vs. congruent). This shows that, despite an across-the-board decrease in reading times as proficiency/experience increased, patterns of performance for idioms versus controls for nonnative English speakers showed no differences according to relative L2 proficiency. It should be noted, however, that exploring individual variation according to proficiency was not a primary aim of this study and was therefore not manipulated. In fact, care was taken to ensure that the nonnative participants had a comparable level of English proficiency. To specifically investigate the influence of proficiency on idiom processing, it would be necessary to test distinct groups of participants at different levels (e.g., English as a foreign language vs. English as a second language learners, like in Yamashita & Jiang, 2010).

DISCUSSION

The Swedish participants in this study showed a consistent advantage when reading idioms compared to literal control phrases. This was true for L2-only idioms, idioms that exist in both the L1 and L2, and L1-only idioms, which by definition should not be familiar in their translated forms. In all conditions, late measures (phrase-level total reading time and regression path durations) confirm that nonnative English speakers had no difficulty understanding the meaning of these phrases and in general spent less time on the idioms than on the literal phrases (when length and single-word frequency were controlled for). This was also partially supported in early measures for the final words (likelihood of skipping), in which Swedish and congruent items but not English items showed an advantage. We interpret

https://doi.org/10.1017/S0272263115000492 Published online by Cambridge University Press

this as evidence that these known combinations were being automatically triggered in such a way that lexical access for the final word was significantly quicker. For English-only idioms, despite the relative ease with which they were understood, no such boost was observed, suggesting that the lexical combinations were not as well entrenched in the mental lexicon, even though the figurative meanings were accessible. English native speakers performed exactly as predicted on English idioms, showing facilitation for the form (via early measures) and meaning integration (via late measures) of idioms compared to matched literal phrases. However, when faced with unfamiliar idioms (translated Swedish forms), they showed considerable disruption in all late measures, suggesting that they had to spend more time reading the idioms in an attempt to work out and integrate their meaning.

L2 Processing of Formulaic Language

The implications for bilingual processing of formulaic language are extremely interesting. The nonnative participants in this study were all at a very high level of proficiency, with only a very small minority (3/24 participants) showing estimated vocabulary sizes of less than 7,000 words. Importantly, the reading patterns indicate that high-proficiency participants are able to easily activate the figurative meanings of English idioms. Clearly, then, there is nothing fundamentally stopping L2 speakers from instantiating idioms in the mental lexicon in a way that enables them to process them quickly, in the same way as native speakers. Equally clear, however, is that the exposure and level of proficiency necessary for this to happen is high: even for the advanced learners in this study, the advantage was modest and was not really evident in the most automatic lexical access measures (skipping rates and early measures for the final words) for the English-only idioms. Although the effects for English-only (L2) idioms were not as clear-cut as for the English native speakers processing familiar phrases in their L1, there is evidence that, through exposure and experience, idioms—despite their noncompositional nature—can become easier to process for nonnative speakers.

Formulaic Transfer from the L1

Of potentially greater interest is the clear finding that nonnative English speakers did treat L1 idioms like formulaic units when these were encountered in the L2, in the sense that they showed the same kind of speeded

https://doi.org/10.1017/S0272263115000492 Published online by Cambridge University Press

processing observed when native speakers encounter idioms. This was true for congruent items, which conceivably could have been encountered in English as well as in Swedish, but also for the Swedish-only items, for which this cannot be the case. The only source of knowledge about these configurations is that the same words go together in the L1, and it is highly unlikely that any of these combinations would ever have been encountered (with the same idiomatic meaning) by the Swedish participants in English. Importantly, despite the unfamiliar form of these translations, there is a clear advantage for idioms versus literal controls, especially in terms of the ease with which these were understood in the overall context of the sentence. Nonnative English speakers had no difficulty in integrating the phrase-level meaning of these items (as shown via the late measures) and showed some evidence that the expected word was being activated, even in the "wrong" language (higher skipping rates for idiom final words in the Swedish and congruent conditions). Importantly, this was the case despite the fact that no biasing context was provided, and despite the fact that all idioms were short—hence there was no unequivocal recognition point until the final word had been read.

There is also no clear evidence that congruency has any additional facilitative effect over and above those items that exist only in the L1. Titone et al. (2015) suggested that their results—less disruption during code-switching of idioms when the items were congruent—provide evidence for the representation of holistic idiom forms in both languages. Our study would dispute this, because there is no evidence that congruent items were treated any differently from Swedish-only items. L1 knowledge appears to be the main driver of this effect, irrespective of whether the item is also known in the L2. The effect of relative familiarity is important here. For both Swedish-only and congruent items, increased familiarity with the Swedish version of the phrase showed a facilitatory effect for idioms in late measures. Thus, items that were better known in the L1 were more easily understood when encountered in the L2. Crucially, the congruent items showed no evidence that familiarity with the specific English form had any effect, which implicates L1 knowledge over and above direct experience in the L2 in how these items were processed. In other words, whether or not these items were also known in the L2, it was the familiarity with the L1 version that determined how easily they were understood. In the case of English-only idioms, for which no existing L1 knowledge exists to aid either the form or meaning of the idioms, experience directly in the L2 shows a clear facilitatory effect. This again suggests that nonnative speakers can develop nativelike formulaic performance in the L2, in line with various studies that have shown this to be the case at high levels of proficiency (Gyllstad & Wolter, 2016; Isobe, 2011; Yamashita & Jiang, 2010; Yeganehjoo & Thai, 2012).

https://doi.org/10.1017/S0272263115000492 Published online by Cambridge University Press

Activation of L1 Lexical Networks

On the question of why L1 knowledge should show such a strong influence, an increasing body of evidence suggests that when bilinguals process language in their L2, they demonstrate ballistic activation (Phillips, Segalowitz, O'Brien, & Yamasaki, 2004). That is, they obligatorily activate the equivalent words in their native language (Wu, Cristino, Leek, & Thierry, 2013; Wu & Thierry, 2010; T. Zhang, van Heuven, & Conklin, 2011). Assuming that this is the case, we can speculate why both congruent and translated forms might show privileged processing in the same way that we see for native speakers encountering L1 forms. Reading the first word of an idiom will automatically trigger the L1 equivalent (e.g., *break* = *bryta*). If we assume that idioms do have a holistic idiom form—either as part of something akin to a "superlemma" (Sprenger et al., 2006) or as part of a duplicate lexical entry encompassing a larger unit (Wray, 2002)—then any known L1 idioms will also be activated by this, leading to activation of the whole unit (*bryta isen*), which in turn will provide a boost in lexical access to the expected word (*isen/ice*). For the control phrase, encountering *crack* will also trigger the L1 equivalent word (*knäcka*), but because *knäcka isen* is not an idiom in either language, no whole-form entry or association between the two words can exist; therefore, there is no reason for *isen/ice* to be activated over and above any other plausible continuation. Under this view, both Swedish-only and congruent items should activate L1 equivalents, leading to facilitation. English-only idioms have no L1 equivalents, but experience in the L2 may have developed entries for at least some idioms (presumably the most frequent/common ones), leading to the modest level of facilitation seen in our results, and the clear effect of specific L2 familiarity in this condition. We should reiterate, however, that our results cannot confirm or disprove such an account, given that the speeded processing seen throughout cannot necessarily be taken as an indicator of holistic processing.

A lemma-based explanation is conceptually very similar. In line with the view put forward by Jiang (2000), Ueno (2009), and Wolter and Yamashita (2014), a learned L2 form may in the first instance link to lemma-level information from the L1. A second stage may occur whereby this lemma is copied to the L2 to give a dedicated L2 lexeme-lemma pairing, but the underlying information still fundamentally reflects the L1. Lexical networks and associations between words may therefore hold in both the L1 and the L2, because the same connections are assumed to underlie the different word forms. Alternatively, lemma-level information may be language nonspecific, with information such as semantic and associative networks being tied to the conceptual values of words rather than a language-specific form (akin to the revised hierarchical model

https://doi.org/10.1017/S0272263115000492 Published online by Cambridge University Press

of Kroll & Stewart, 1994, whereby specific L1/L2 forms link to a shared conceptual store). This may also explain why, for congruent items, we see an effect of well-established L1 familiarity over and above any effect of specific L2 familiarity, as this is likely to be much more strongly established and linked to the underlying concept/lemma. One way to test this might be to perform this study in reverse by translating the English items into Swedish to see how L1 Swedes process them. If lemmas are language nonspecific, then we should see some level of facilitation for Swedish-English speakers, whereas Swedish participants with no knowledge of English should show disruption, as seen in the present study for English native speakers reading translated Swedish forms.

In summary, our results show clear support for L1 influence on the processing of idioms by advanced-proficiency nonnative speakers. Importantly, our study suggests that this knowledge—in addition to being evident in offline tasks, as shown in previous research (Bulut & Çelik-Yazici, 2004; Charteris-Black, 2002; Irujo, 1986; Laufer, 2000; Liontas, 2001)—is used in an online fashion, facilitating lexical access and semantic integration for known combinations from the very earliest stages of processing, and leading to faster processing in the same way as we see for native speakers. The fact that this is true whether or not the combination also exists in the L2 is crucial because it prioritizes L1 knowledge directly, rather than fitting a confirmatory account whereby L2 idioms have been encountered and mentally registered as transferrable in the minds of individual learners, or in which congruent idioms are dually represented in both the L1 and L2 lexicons.

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit http://dx.doi.org/S0272263115000492.

Received 6 February 2015
Accepted 17 November 2015
Final Version Received 19 November 2015

NOTES

1. Williams and Cheung tested semantic priming from L3 (French) to L1 (Chinese) via English (L2), which was the language of instruction. They used French prime words (e.g., *chaise*) and Chinese target words (e.g., 書桌 [*desk*]), on the assumption that because English had been the language of instruction, no episodic memory connections could exist between the French and Chinese forms, and hence any priming should be the effect of direct semantic connections. For the sake of simplicity, we have presented only the English-French forms to demonstrate the priming effects that were/were not observed.

2. The shortened version was used, for practical reasons, because a full-length vocabulary test in addition to the eye-tracking study and collection of subjective rating data (detailed later in this section) could have led to fatigue and could have influenced responses (Bachman & Palmer, 1996).

3. This reliability coefficient was reached based on data for 10 of the 20 items, because 10 items had zero variance and therefore did not contribute to the scale. Considering this, an alpha of close to .80 must be considered satisfactory for this very short vocabulary test.

4. Comprehension scores based on proportion of correct answers were very high among nonnative participants, with a mean of 92% (*SD* = 4.8, range = 83–100). This supports our assumption that the learners in this study were of a fairly advanced level. In particular, it is worth noting that the three individuals who scored lowest in the modified vocabulary size test did not differ markedly from this group mean (scores of 90%, 85%, and 95% respectively).

REFERENCES

Abel, B. (2003). English idioms in the first language and second language lexicon: A dual representation approach. *Second Language Research*, *19*, 329–358.

Alderson, C. (2005). *Diagnosing foreign language proficiency: The interface between learning and assessment*. London: Continuum.

Altarriba, J., Kroll, J., Sholl, A., & Rayner, K. (1996). The influence of lexical and conceptual constraints on reading mixed-language sentences: Evidence from eye fixations and naming times. *Memory and Cognition*, *24*, 477–492.

Arcara, G., Lacaita, G., Mattaloni, E., Passarini, L., Mondini, S., Beninca, P., & Semenza, C. (2012). Is "hit and run" a single word? The processing of irreversible binomials in neglect dyslexia. *Frontiers in Psychology*, *3*, 1–11.

Bachman, L., & Palmer, A. (1996). *Language Testing in Practice: Designing and Developing Useful Language Tests*. Oxford: Oxford University Press.

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

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

Bates, D. M., Maechler, M., Bolker, B., Walker, S., Christensen, R., Singmann, H., & Dai, B. (2014). lme4: Linear mixed-effects models using Eigen and S4 (Version 1.0–7) [Computer software].

Blais, M.-J., & Gonnerman, L. M. (2013). Explicit and implicit semantic processing of verb–particle constructions by French–English bilinguals. *Bilingualism: Language and Cognition*, *16*, 829–846.

Boers, F., & Lindstromberg, S. (2012). Experimental and intervention studies on formulaic sequences in a second language. *Annual Review of Applied Linguistics*, *32*, 83–110.

Bulut, T., & Çelik-Yazici, I. (2004). Idioms processing in L2: Through rose-colored glasses. *The Reading Matrix*, *4*, 105–116.

Bybee, J. L. (2006). From usage to grammar: The mind's response to repetition. *Language*, *82*, 711–733.

Cacciari, C., & Tabossi, P. (1988). The comprehension of idioms. *Journal of Memory and Language*, *27*, 668–683.

Carrol, G., & Conklin, K. (2014a). Eye-tracking multi-word units: Some methodological questions. *Journal of Eye Movement Research*, *7*, 1–11.

Carrol, G., & Conklin, K. (2014b). Getting your wires crossed: Evidence for fast processing of L1 idioms in an L2. *Bilingualism: Language and Cognition*, *17*, 784–797.

Carrol, G., & Conklin, K. (2015, April 20). Cross language priming extends to formulaic units: Evidence from eye-tracking suggests that this idea "has legs." *Bilingualism: Language and Cognition*. http://dx.doi.org/10.1017/S1366728915000103.

Charteris-Black, J. (2002). Second language figurative proficiency: A comparative study of Malay and English. *Applied Linguistics*, *23*, 104–133.

Cieślicka, A. B. (2006). Literal salience in on-line processing of idiomatic expressions by second language learners. *Second Language Research*, *22*, 115–144.

Cieślicka, A. B. (2013). Do nonnative language speakers chew the fat and spill the beans with different brain hemispheres? Investigating idiom decomposability with the divided visual field paradigm. *Journal of Psycholinguistic Research*, *42*, 475–503.
Cieślicka, A. B., & Heredia, R. R. (2011). Hemispheric asymmetries in processing L1 and L2 idioms: Effects of salience and context. *Brain and Language*, *116*, 136–150.
Cieślicka, A. B., Heredia, R. R., & Olivares, M. (2014). It's all in the eyes: How language dominance, salience, and context affect eye movements during idiomatic language processing. In M. Pawlak & L. Aronin (Eds.), *Essential Topics in Applied Linguistics and Multilingualism* (pp. 21–41). Switzerland: Springer International.
Clark, H. H. (1973). The language as-fixed-effect fallacy. *Journal of Verbal Learning and Verbal Behavior*, *12*, 335–359.
Clifton, C., Staub, A. & Rayner, K. (2007). Eye movements in reading words and sentences. In R. Van Gompel, M. Fisher, W. Murray, & R. L. Hill (Eds.), *Eye movement research: A window on mind and brain* (pp. 341–372). Oxford: Elsevier.
Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? *Applied Linguistics*, *29*, 72–89.
Cutter, M. G., Drieghe, D., & Liversedge, S. P. (2014). Preview benefit in English spaced compounds. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *40*, 1778–1786.
Davies, M. (2004). *BYU-BNC* (Based on the British National Corpus from Oxford University Press) [Online resource]. Retrieved from http://corpus.byu.edu/bnc/.
De Groot, A., & Nas, G. (1991). Lexical representation of cognates and noncognates in compound bilinguals. *Journal of Memory and Language*, *30*, 90–123.
De Rosario-Martinez, H. (2013). Phia: Post-hoc interaction analysis (R package).
Durrant, P., & Schmitt, N. (2010). Adult learners' retention of collocations from exposure. *Second Language Research*, *26*, 163–188.
Gyllstad, H. (2007). *Testing English collocations: Developing tests for use with advanced Swedish learners* (Doctoral dissertation). Lund: Lund University. Retrieved from http://lup.lub.lu.se/record/599011/file/2172422.pdf
Gyllstad, H. (2012, May 31–June 3). *Validating the vocabulary size test—a classical test theory approach*. Poster presented at the 9th EALTA conference, Innsbruck, Austria. Retrieved from http://www.ealta.eu.org/conference/2012/posters/Gyllstad.pdf.
Gyllstad, H., & Wolter, B. (2016). Collocational processing in light of a phraseological continuum model: Does semantic transparency matter? *Language Learning*, *66*.
Hargevik, S., & Ljung, M. (1989). Say when *och 1.284 andra engelska idiom* [Say *when* and 1,284 other English idioms]. Solna: Esselte studium.
Holsinger, E. (2013). Representing idioms: Syntactic and contextual effects on idiom processing. *Language and Speech*, *56*, 373–394.
Hübinette, L., & Odenstedt, B. (1988). *Ord och inga visor—2000 svenska idiom i engelsk översättning* [Straight talking—2000 Swedish idioms with English translations]. Lund: Studentlitteratur.
Inhoff, A. (1984). Two stages of word processing during eye fixations in the reading of prose. *Journal of Verbal Learning and Verbal Behavior*, *23*, 612–624.
Irujo, S. (1986). *Don't put your leg in your mouth: Transfer in the acquisition of idioms in a second language*. *TESOL Quarterly*, *20*, 287–304.
Isobe, Y. (2011). Representation and processing of formulaic sequences in L2 mental lexicon: How do Japanese EFL learners process multi-word expressions? *JACET Kansai Journal*, *13*, 38–49.
Jackendoff, R. (1995). The boundaries of the lexicon. In M. Everaert, E. van der Linden, A. Schenk, & R. Schreuder (Eds.), *Idioms: Structural and psychological perspectives* (pp. 133–166). Hillsdale, NJ: Erlbaum.
Jiang, N. (2000). Lexical representation and development in a second language. *Applied Linguistics*, *21*, 47–77.
Jiang, N., & Nekrasova, T. (2007). The processing of formulaic sequences by second language speakers. *The Modern Language Journal*, *91*, 433–445.
Jolsvai, H., McCauley, S. M., & Christiansen, M. H. (2013). Meaning overrides frequency in idiomatic and compositional multiword chunks. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *35th Annual Conference of the Cognitive Science Society* (pp. 392–697). Austin, TX: Cognitive Science Society.

Kecskés, I. (2000). A cognitive-pragmatic approach to situation-bound utterances. *Journal of Pragmatics*, *32*, 605–625.
Kellerman, E. (1977). Towards a characterization of the strategy of transfer in second language learning. *Interlanguage Studies Bulletin*, *2*, 58–145.
Kellerman, E. (1986). An eye for an eye: Crosslinguistic constraints on the development of the L2 lexicon. In E. Kellerman & M. Sharwood Smith (Eds.), *Crosslinguistic influence in second language acquisition* (pp. 35–48). Oxford: Pergamon Press.
Kroll, J., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, *33*, 149–174.
Kuznetsova, A., Brockhoff, P., & Christensen, R. (2014). lmerTest: Tests in Linear Mixed Effects Models (Version 2.0–11) [Computer software].
Laufer, B. (2000). Avoidance of idioms in second language: The effect of L1-L2 degree of similarity. *Studia Linguistica*, *54*, 186–196.
Laufer, B., & Waldman, T. (2011). Verb-noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning*, *61*, 647–672.
Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
Libben, M., & Titone, D. (2008). The multidetermined nature of idiom processing. *Memory and Cognition*, *36*, 1103–1121.
Liontas, J. (2001). That's all Greek to me! The comprehension and interpretation of modern Greek phrasal idioms. *The Reading Matrix*, *1*, 1–32.
Matlock, T., & Heredia, R. R. (2002). Understanding phrasal verbs in monolinguals and bilinguals. In R. R. Heredia & J. Altarriba (Eds.), *Bilingual sentence processing* (pp. 251–274). Amsterdam: Elsevier.
McGlone, M., Glucksberg, S., & Cacciari, C. (1994). Semantic productivity and idiom comprehension. *Discourse Processes*, *17*, 167–190.
Meara, P., & Jones, G. (1988). Vocabulary size as a placement indicator. In P. Grunwell (Ed.), *Applied linguistics in society* (pp. 80–87). London: CILT.
Milton, J. (2010). The development of vocabulary breadth across the CEFR levels: A common basis for the elaboration of language syllabuses, curriculum guidelines, examinations, and textbooks across Europe. In I. Bartning, M. Martin, & I. Vedder (Eds.), *Communicative proficiency and linguistic development: Intersections between SLA and language testing research* (pp. 211–232). Eurosla Monograph Series I. Retrieved from http://eurosla.org/monographs/EM01/EM01tot.pdf
Moon, R. (1998). *Fixed expressions and idioms in English: A corpus-based approach*. Oxford: Clarendon Press.
Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review*, *63*, 59–82.
Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, *31*, 9–13.
Paterson, K., Liversedge, S., & Underwood, G. (1999). The influence of focus operators on syntactic processing of short relative clause sentences. *The Quarterly Journal of Experimental Psychology*, *52A*, 717–737.
Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 191–225). London: Longman.
Phillips, N. A., Segalowitz, N., O'Brien, I., & Yamasaki, N. (2004). Semantic priming in a first and second language: Evidence from reaction time variability and event-related brain potentials. *Journal of Neurolinguistics*, *17*, 237–262.
Pickering, M. J., Frisson, S., McElree, B., & Traxler, M. (2004). Eye movements and semantic composition. In M. Carreiras & C. Clifton (Eds.), *On-line study of sentence comprehension: Eyetracking, ERPs and beyond* (pp. 33–50). New York: Psychology Press.
R Development Core Team. (2014). R: A Language and Environment for Statistical Computing (Version 3.1.2) [Computer software]. Vienna, Austria: R Foundation for Statistical Computing.
Rommers, J., Dijkstra, T., & Bastiaansen, M. (2013). Context-dependent semantic processing in the human brain: Evidence from idiom comprehension. *Journal of Cognitive Neuroscience*, *25*, 762–776.
Schweigert, W. (1986). The comprehension of familiar and less familiar idioms. *Journal of Psycholinguistic Research*, *15*, 33–45.

Siyanova-Chanturia, A. (2015). On the "holistic" nature of formulaic language. *Corpus Linguistics and Linguistic Theory*. doi:10.1515/cllt-2014-0016
Siyanova-Chanturia, A., Conklin, K., & Schmitt, N. (2011). Adding more fuel to the fire: An eye-tracking study of idiom processing by native and non-native speakers. *Second Language Research*, *27*, 251–272.
Smolka, E., Rabanus, S., & Rösler, F. (2007). Processing verbs in German idioms: Evidence against the configuration hypothesis. *Metaphor and Symbol*, *22*, 213–231.
Sprenger, S., Levelt, W., & Kempen, G. (2006). Lexical access during the production of idiomatic phrases. *Journal of Memory and Language*, *54*, 161–184.
Staub, A., & Rayner, K. (2007). Eye movements and on-line comprehension processes. In M. G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 327–342). Oxford: Oxford University Press.
Swinney, D., & Cutler, A. (1979). The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behaviour*, *18*, 523–534.
Tabossi, P., Fanari, R., & Wolf, K. (2009). Why are idioms recognized fast? *Memory and Cognition*, *37*, 529–540.
Titone, D., Columbus, G., Whitford, V., Mercier, J., & Libben, M. (2015). Contrasting bilingual and monolingual idiom processing. In R. R. Heredia & A. B. Cieślicka (Eds.), *Bilingual figurative language processing* (pp. 171–207). Cambridge: Cambridge University Press.
Titone, D., & Connine, C. (1999). On the compositional and noncompositional nature of idiomatic expressions. *Journal of Pragmatics*, *31*, 1655–1674.
Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
Ueno, T. (2009). *An investigation of the relationship between the development of bilingual semantic organisation and interactive connectivity across languages* (Unpublished doctoral dissertation). Trinity College, Dublin.
Warren, H. (1994). *Oxford learner's dictionary of English idioms*. Oxford: Oxford University Press.
Williams, J. (1994). The relationship between word meanings in the first and second language: Evidence for a common, but restricted, semantic code. *European Journal of Cognitive Psychology*, *6*, 195–220.
Williams, J., & Cheung, A. (2011). Using priming to explore early word learning. In P. Trofimovich & K. McDonough (Eds.), *Applying priming methods to L2 learning, teaching and research* (pp. 73–104). Amsterdam: John Benjamin's.
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: A comparison of congruent and incongruent collocations. *Studies in Second Language Acquisition*, *35*, 451–482.
Wolter, B., & Yamashita, J. (2014). Processing collocations in a second language: A case of first language activation? *Applied Psycholinguistics*, 1–29.
Wray, A. (2002). *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.
Wray, A. (2012). What do we (think we) know about formulaic language? An evaluation of the current state of play. *Annual Review of Applied Linguistics*, *32*, 231–254.
Wu, Y. J., Cristino, F., Leek, C., & Thierry, G. (2013). Non-selective lexical access in bilinguals is spontaneous and independent of input monitoring: Evidence from eye tracking. *Cognition*, *129*, 418–425.
Wu, Y. J., & Thierry, G. (2010). Chinese-English bilinguals reading English hear Chinese. *The Journal of Neuroscience*, *30*, 7646–7651.
Wulff, S. (2008). *Rethinking idiomaticity—a usage-based approach*. London/New York: Continuum.
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.
Yeganehjoo, M., & Thai, Y. N. (2012). Lexical access in production of idioms by proficient L2 learners. *The Southeast Asian Journal of English Language Studies*, *18*, 87–104.
Zhang, T., van Heuven, W. J., & Conklin, K. (2011). Fast automatic translation and morphological decomposition in Chinese-English bilinguals. *Psychological Science*, *22*, 1237–1242.
Zhang, X. (2013). The I don't know option in the vocabulary size test. *TESOL Quarterly*, *47*, 790–811.

APPENDIX A

EXPERIMENTAL ITEMS

English		Congruent		Swedish	
Idiom	Control	Idiom	Control	Idiom	Control
Bite the bullet	Grab the bullet	Bear fruit	Grow fruit	Born in the hall	Left in the hall
Blow a fuse	Need a fuse	Bend the rules	Read the rules	Chew foam	Use foam
Break the bank	Hurt the bank	Bide your time	Use your time	Come on shame	Focus on shame
Chew the fat	Use the fat	Bite your tongue	Burn your tongue	Confess colour	Change colour
Clear the decks	Wash the decks	Break the ice	Crack the ice	Cow on the ice	Game on the ice
Cook the books	Check the books	Break the silence	End the silence	Crawl to the cross	Pray to the cross
Cross your fingers	Mind your fingers	Burn your boats	Lose your boats	Cream on the mash	Sauce on the mash
Cut your losses	Count your losses	Bury the hatchet	Find the hatchet	Cry rivers	Use rivers
Drop the ball	Miss the ball	Call your bluff	Match your bluff	Get the kick	Miss the kick
Face the music	Play the music	Clear the air	Check the air	Give him the basket	Sell him the basket
Find your feet	Hurt your feet	Draw a blank	Leave a blank	Give the iron	Sell the iron
Foot the bill	Read the bill	Drown your sorrows	Express your sorrows	Hang lip	Give lip
Hit the roof	Fix the roof	Eat your words	Know your words	Hard bandages	New bandages
Hold the fort	Take the fort	Fall from grace	Slip from grace	Harvest victims	Collect victims
Hold your horses	Lead your horses	Gain ground	Clear ground	Hold box	Never box
Jump the gun	Take the gun	Gather dust	Produce dust	Hot on the porridge	Keen on the porridge
Keep your head	Mind your head	Have a point	Deserve a point	Lose the suction	Apply the suction
Kick the bucket	Drop the bucket	Keep the peace	Like the peace	Make a painting	Buy a painting
Know the ropes	Bring the ropes	Learn your lesson	Finish your lesson	Neck over head	Back over head

Continued

https://doi.org/10.1017/S0272263115000492 Published online by Cambridge University Press

Appendix A. Continued

English		Congruent		Swedish	
Idiom	Control	Idiom	Control	Idiom	Control
Lose your marbles	Count your marbles	Lick your wounds	Dress your wounds	Play monkey	Taste monkey
Make a scene	Paint a scene	Lose the thread	Pull the thread	Pull logs	Cut logs
Mark your words	Hear your words	Lose your head	Hurt your head	Shoulder his coat	Carry his coat
Pick a fight	Have a fight	Meet your maker	Call your maker	Similar as berries	Tasty as berries
Pick your brains	Use your brains	Meet your match	Win your match	Sit inside	Stay inside
Pop the question	Shout the question	Miss the point	Pass the point	Smell cat	Hear cat
Pull your leg	Grab your leg	Pass the time	Use the time	Stand on the nose	Focus on the nose
Pull your weight	Control your weight	Play with fire	Cook with fire	Step in the piano	Load in the piano
Push your luck	Make your luck	Show your face	Paint your face	Suck on the frames	Grow on the frames
Risk your neck	Hurt your neck	Steal the show	Like the show	Take battle	Risk battle
Rock the boat	Crash the boat	Stretch your legs	Move your legs	Take it piano	Be it piano
Save the day	Ruin the day	Swallow your pride	Regain your pride	Take screw	Need screw
Smell a rat	Hear a rat	Sweeten the pill	Swallow the pill	The red thread	The main thread
Spill the beans	Drop the beans	Take a joke	Tell a joke	The whole ballet	The new ballet
Stand your ground	Keep your ground	Take shape	Lose shape	Throw water	Find water
Take your pick	Make your pick	Tighten your belt	Change your belt	Toil dog	Eat dog
Toe the line	Mark the line	Tread water	Lose water	Turn the steak	Cook the steak
Turn the tables	Move the tables	Try your luck	Fix your luck	Under the ice	Into the ice
Waste your breath	Lose your breath	Turn the screw	Find the screw	Understand the gallop	Hear the gallop
Watch your step	Clean your step	Wait your turn	Miss your turn	Walk away	Move away
Weather the storm	Monitor the storm	Watch the clock	Mend the clock	Weak comfort	Small comfort

https://doi.org/10.1017/S0272263115000492 Published online by Cambridge University Press

APPENDIX B

EXAMPLES OF CONTEXT SENTENCES

English idiom *spill the beans* (meaning: reveal a secret) vs. control phrase *drop the beans*:

It was hard not to *spill the beans* when I heard such a juicy piece of gossip.
It was hard not to *drop the beans* after I cut myself when I was opening the can.

Congruent idiom *play with fire* (meaning: do something risky) vs. control phrase *cook with fire*:

My friend's been *playing with fire* and it was always likely to get him into trouble.
My friend's been *cooking with fire* and it's given the meat a really nice smoky flavour.

Swedish idiom *shoulder his coat* (meaning: live up to his success) vs. control phrase *carry his coat*:

I'm not sure I can *shoulder his coat* because he's had so much success in the past.
I'm not sure I can *carry his coat* because I have all of my own things and my hands are full.

https://doi.org/10.1017/S0272263115000492 Published online by Cambridge University Press