Applied Psycholinguistics (2026), **47**, e64, 1–31
doi:10.1017/S0142716426100939

CAMBRIDGE UNIVERSITY PRESS

ORIGINAL ARTICLE

The uneven advantage of multiword sequences: How idioms and collocations are processed in L1 and L2

Wanyin Li[1], Bene Bassetti[1,2] and Steven Frisson[3]

[1]School of Education, University of Birmingham, UK, [2]Department of Education and Humanities, University of Modena and Reggio Emilia, Italy and [3]School of Psychology, University of Birmingham, UK
Corresponding author: Wanyin Li; Email: wanyin.li.words@gmail.com

(Received 12 August 2025 UTC; revised 25 June 2026 UTC; accepted 2 July 2026 UTC)

Abstract

Research has shown a processing advantage for multiword sequences (MWS) in both first language (L1) and second language (L2) speakers over novel word combinations. However, whether this advantage holds for different types of MWS for both groups is not known. We used a self-paced reading paradigm combined with a stop-making-sense task to investigate the processing of idioms used figuratively (e.g., "gray area"), idioms used literally, collocations (e.g., "vast area"), and novel combinations (e.g., "rare area"), in both L1 and L2 speakers. In line with previous findings, both L1 and L2 speakers read collocations faster than novel word combinations. Only L1 speakers showed a processing advantage for idioms used figuratively over collocations after controlling for distributional factors (e.g., phrase-level frequency, familiarity, and contextual predictability), suggesting that semantic conventionality facilitates direct retrieval of idiomatic meanings. L2 speakers with lower proficiency did not show an advantage for idioms used figuratively but exhibited faster reading for idioms used literally over idioms used figuratively. While the meaning dominance of idioms did not exert an effect, greater idiom decomposability facilitated processing. Together, these findings indicate that, while frequency effects broadly support MWS processing, L2 speakers rely on compositional analysis in idiom comprehension more than L1 speakers.

Keywords: Adult second language acquisition; lexical processing; reading; sentence processing

Introduction

A multiword sequence (MWS; Christiansen & Arnon, 2017), also referred to as formulaic language (Carrol & Conklin, 2020) or a multiword expression (Kallens & Christiansen, 2022, Sag et al., 2002) is a highly familiar frequently occurring word combination that serves as fundamental building blocks of language (Arnon et al., 2017; Arnon & Snider, 2010). MWSs encompass different

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

Downloaded from https://www.cambridge.org/core. 26 Oct 2026 at 14:05:28, subject to the Cambridge Core terms of use.

linguistic forms, such as idioms (*shoot the breeze*), collocations (*fatal mistake*), binomials (*bride and groom*), lexical bundles (*on the other hand*), and other expressions (see Wray, 2002). MWSs are highly pervasive in both spoken and written language; for example, Erman and Warren (2000) estimated that about 55% of a text contains formulaic expressions of some sort.

MWSs are generally associated with a processing advantage, being processed more quickly than word sequences that do not frequently co-occur (e.g., Li et al., 2025). This processing advantage challenges the traditional "words-and-rules" view (Pinker, 1999) of language use and processing (for discussion, see Kallens & Christiansen, 2022). Rather than solely relying on the incremental combination of individual words within a syntactic frame, MWSs appear to benefit from language users' sensitivity to recurrent occurring patterns, giving rise to a degree of unit-like representation and processing.

Historically, most research of MWSs was confined to the processing of idioms (Siyanova-Chanturia & Martinez, 2015). While MWSs research nowadays also includes different types of expressions, direct comparisons between different types are scarce and further complicated by MWSs varying in several important aspects, such as the degree of conventionalization (i.e., the extent to which the form configuration or meaning is static), semantic decomposability/transparency (i.e., the extent to which individual components of a MWS independently contribute to the overall meaning), and meaning dominance (i.e., the extent that MWSs are used in their derived or compositional meaning). The present study directly compares two common types of MWS: collocations (e.g., *vast area*) and ambiguous idioms (i.e., word sequences that can have both a literal and figurative interpretation, e.g., *gray area*). Moreover, MWSs are important in second language acquisition; being able to competently use the right MWS in the right context is indicative of L2 proficiency (Pawley & Syder, 1983). Given that different types of MWS pose different challenges for L2 users (Yi & Zhong, 2024), we will compare the processing of both types in both L1 and L2 (Chinese-English) language users.

To our knowledge, only one study has investigated to what extent distributional properties, such as frequency and contextual predictability, affect different MWSs. Carrol and Conklin (2020) recorded eye movements while participants read sentences containing idioms (*spill the beans*), collocations (*abject poverty*), binomials (*king and queen*), or matched novel phrases (*drop the beans; total poverty; prince and queen*). They examined whether the processing advantage typically observed for MWSs over novel phrases was consistent across different MWS types, and how far this advantage could be explained by general (e.g., frequency) and/or type-specific (e.g., degree of decomposability) properties. They found a consistent advantage for MWSs over novel phrases and argued that this effect was mainly driven by frequency and predictability, i.e., by general properties. They concluded that the processing advantage for MWS is largely driven by "aspects of conventionalization (frequency [. . .] and predictability [. . .] based on co-occurrence probabilities)" (p. 116).

The present study expands Carrol and Conklin's (2020) research in several important ways. First, the different MWSs, due to their different configurations, could not be compared directly, which our design did allow. Second, in our study, distributional variables such as frequency, familiarity and predictability, were

Downloaded from https://www.cambridge.org/core. 26 Oct 2026 at 14:05:28, subject to the Cambridge Core terms of use.

carefully controlled between all MWSs. Third, we included idioms in both literal and figurative contexts, enabling us to exploit their inherent ambiguity to investigate their processing. Fourth, while Carrol and Conklin only tested native participants, we tested both L1 and L2 participants and compared the processing profiles between the two groups, providing a broader perspective on how MWSs are processed across language backgrounds.

Collocation processing in the first and second language

Previous studies suggest that collocations (e.g., *provide information, fatal mistak*e) are processed more quickly than novel word combinations that infrequently co-occur (e.g., *compare information, extreme mistake*) by both L1 and L2 speakers. This processing advantage has been demonstrated using different paradigms, including priming tasks (Durrant & Doherty, 2010; Vilkaitė-Lozdienė & Conklin, 2021), phrasal judgment tasks (e.g., Siyanova & Schmitt, 2008; Wolter & Gyllstad, 2011, 2013; Wolter & Yamashita, 2015, 2018), and eye-tracking during reading (Li et al., 2021; Li et al., 2025; Sonbul, 2015; Vilkaitė & Schmitt, 2019; Vilkaitė, 2016).

Furthermore, both L1 and L2 speakers are sensitive not only to the frequency of individual words within collocations but also to the frequency of the collocation as a whole (Öksüz et al., 2021, 2024). Using an acceptability judgment task, Öksüz et al. (2021) found that both noun frequency and collocational frequency were associated with faster responses in both L1 and L2 speakers. Öksüz et al. (2021) also found that higher collocational frequency reduced reliance on word frequency, indicating a shift toward holistic processing for frequent collocations. This suggests that frequent collocations become more entrenched and lexicalized in memory, consistent with usage-based language models (Bybee, 1998; Goldberg, 2006; Tomasello, 2003). However, given that both groups of speakers remained sensitive to word-level frequency, even in frequent collocations, lexicalization does not mean collocations are processed solely as chunks without compositional analysis.

Studies reporting a collocation advantage (Carrol & Conklin, 2020; Durrant & Doherty, 2010; Li et al., 2021; Li et al., 2025; Öksüz et al., 2021; Siyanova & Schmitt, 2008; Sonbul, 2015; Vilkaitė, 2016) typically adopted a frequency-based approach (Sinclair, 1991) to define collocations. Collocations can be defined using two distinct approaches: the corpus-based frequency method (Sinclair, 1991) and the phraseology-based method (e.g., Howarth, 1998). The frequency method defines collocations as frequently co-occurring word combinations and includes semantically transparent and literal collocations (Carrol & Conklin, 2020). In contrast, the phraseology-based method emphasizes semantic restrictions, where one component may not be used literally or lacks semantic transparency (e.g., *heavy rain*), as described in Howarth's Continuum Model (1998). For the current study, we employed the corpus-based frequency method and selected only transparent and literal collocations. Full details are provided in the materials section.

Idiom superiority effects in the first language

Idioms are characterized as a string of words for which the figurative meaning cannot be derived by computing the individual components literally (Swinney &

Downloaded from https://www.cambridge.org/core. 26 Oct 2026 at 14:05:28, subject to the Cambridge Core terms of use.

Cutler, 1979). Unlike collocations, idioms are perceived as fixed, semantically opaque, and having a conventionalized figurative meaning. A considerable body of literature demonstrates idioms' superiority effect in L1 processing, i.e., faster processing of idioms compared to novel literal word combinations, with evidence coming from lexical decision tasks (e.g., Ortony et al., 1978; Swinney & Culter, 1979; Tabossi et al., 2009), self-paced reading tasks (e.g., Conklin & Schmitt, 2008; Gibbs, 1980), and eye-tracking studies (Carrol & Conklin, 2017, 2020; Siyanova-Chanturia et al., 2011; Underwood et al., 2004).

The well-documented idiom superiority effect in L1 speakers was initially taken as evidence that idioms are represented as non-decomposable strings and retrieved directly from the mental lexicon (e.g., Bobrow & Bell, 1973; Swinney & Culter, 1979). However, subsequent research showed that this effect is primarily driven by the high familiarity among L1 speakers, facilitating faster processing than novel word combinations (e.g., Carrol & Conklin, 2017, 2020; Tabossi et al., 2009). Furthermore, studies showed that idiom processing involves both literal computation, which is considered obligatory, and figurative retrieval (e.g., Cacciari & Corradini, 2015; Cacciari & Tabossi, 1988; Cutting & Bock, 1997; Titone et al., 2019). Hence, idioms are characterized as simultaneously compositional and non-decomposable (Kuiper et al., 2007), exhibiting features of both words and complex syntactic structures. Like words, idioms may have unitary representations. At the same time, they resemble syntactic structures that undergo compositional analysis (Cutting & Bock, 1997; Sprenger et al., 2006).

As discussed above, there appears to be little inherent to idioms that explains their processing superiority. Like collocations, the processing difference between idioms and novel word combinations can largely be due to idioms being more familiar and better known to language users. However, idioms do not entirely fit in the usage-based approach, as individual idioms are relatively infrequent in language, particularly in corpora (Moon, 1998). This discrepancy leads us to question whether idioms still demonstrate the superiority effect compared to other word combinations when distributional information is controlled for.

The present study therefore separates two constructs. Conventionality refers to how strongly an expression is established in the language community. An expression can be conventionalized as a preferred configuration, as an established meaning, or both. Idioms are treated as strongly semantically conventionalized because they not only have a fixed form but also an established figurative meaning. Although corpus-defined collocations involve preferred co-occurrence patterns, they are generally less fixed in form than idioms, and their meanings are usually literal. The second construct, semantic decomposability/transparency, refers to how easily the phrase-level meaning can be understood from the meanings of the individual words. In this study, this dimension was measured only for idioms, because the collocations were selected to be literal, and, hence, transparent. For example, the idiom *break the ice* has the conventionalized figurative meaning "to make people feel more relaxed in a social situation". However, this meaning cannot be fully understood from the literal meanings of *break* and *ice*, making the expression relatively non-decomposable. The present study therefore asks whether higher semantic conventionality facilitates MWS processing beyond frequency, familiarity, and predictability, and whether differences in semantic decomposability

Downloaded from https://www.cambridge.org/core. 26 Oct 2026 at 14:05:28, subject to the Cambridge Core terms of use.

affect idiom processing. These constructs are distinct from frequency and subjective familiarity, which reflect how often an expression is encountered rather than how its meaning is represented.

Some evidence suggests that semantic meaningfulness can facilitate MWS processing over and above frequency. Using a phrasal decision task, Jolsvai et al. (2020) showed that expressions judged as more meaningful (e.g., *play the field*) were recognized faster than those judged less meaningful (e.g., *because it lets*). In addition, Shi et al. (2023) used self-paced reading and reported a trend for faster reading of figurative MWSs with a figuratively used component (e.g., *build a career*) relative to matched literal collocations (e.g., *build a house*) in L1 speakers. Together, these findings are consistent with hybrid accounts of idiom and MWS processing (Kuiper et al., 2007; Sprenger et al., 2006), in which MWSs with more conventionalised form—meaning mappings are more likely to engage direct access to a unitary meaning representation, reducing reliance on fully literal, word-by-word compositional analysis. However, this has not yet been directly tested and whether conventionality facilitates MWS processing over and above distributional information remains unknown.

Moreover, semantic decomposability (i.e., the extent to which an idiom's components directly relate to the idiomatic meaning) has been suggested to affect idiom processing, as in the Idiom Decomposition Hypothesis (Gibbs et al., 1989; Gibbs & Nayak, 1989). Specifically, it influences the bottom-up incremental parsing or compositional analysis. However, its role remains unclear. Some studies reported a facilitatory effect with decomposable idioms being processed faster than non-decomposable ones (Caillies & Butcher, 2007; Carrol & Segaert, 2024; Gibbs & Nayak, 1989; Hamblin & Gibbs, 1999; Libben & Titone, 2008; Milburn et al., 2021; Titone et al., 2019). In contrast, other studies found no difference (Carrol & Conklin, 2014; Kyriaco et al., 2020; Tabossi et al., 2009) or even the opposite effect (Titone & Libben, 2014), with non-decomposable idioms processed more quickly. In addition, Milburn and Warren (2019) showed that semantic decomposability, operationalized as the degree to which an idiom's literal and figurative meanings overlap, interacts with usage frequency to influence processing. Highly decomposable idioms slow reading when their figurative interpretation is dominant but speed it up when the literal and figurative interpretations of idioms are more balanced. This pattern highlights how facilitatory or inhibitory effects of semantic decomposability depend on which interpretation readers expect.

Idiom processing in the second language

The idiom superiority effect is less consistent in L2 speakers than in L1 speakers, likely due to weaker connections between the idiom's configuration and its idiomatic interpretation. Although some studies (e.g., Conklin & Schmitt, 2008) show L2 speakers read idioms faster than novel word combinations, others did not find an idiom superiority effect in L2 speakers (e.g., Carrol et al., 2016; Carrol & Conklin, 2017; Siyanova-Chanturia et al., 2011). Siyanova-Chanturia et al. (2011) compared ambiguous idioms to novel word combinations within short discourse contexts. The preceding context sentences were biased toward either a literal or a figurative

Downloaded from https://www.cambridge.org/core. 26 Oct 2026 at 14:05:28, subject to the Cambridge Core terms of use.

interpretation of the idiom. L1 speakers read idioms in both literal and figurative interpretations more quickly than novel word combinations. However, L2 speakers did not demonstrate faster processing of the idiomatic interpretations over novel word combinations. Moreover, L2 speakers read the literal interpretations of idioms faster than the figurative ones.

The processing cost that L2 speakers experience with the idiomatic meaning suggests that they rely on a more compositional, bottom-up approach to idiom processing (Beck & Weber, 2016; Cieślicka, 2006; Senaldi & Titone, 2022). This view aligns with several idiom processing models. According to the Dual Process model (van Lancker Sidtis, 2012; see also Wray, 2002), the direct retrieval route relies on distributional properties such as familiarity and exposure frequency (Titone & Connine, 1994). This route allows users to recognize and access idiomatic forms and meanings directly, especially for familiar idioms. The computation route is required for less familiar idioms, involving a more analytical parsing process. For familiar idioms, both routes interact, with the direct route typically prevailing—a pattern common among L1 speakers. For less familiar idioms, the computational analysis dominates over direct retrieval. Given their limited exposure to L2 idioms, L2 speakers are more likely to rely on the computational route, diminishing the idiom superiority effect. According to the Literal Salience model (Cieślicka, 2006), L2 speakers default to a literal analysis of idioms even when they know the figurative meaning and only compute the figurative interpretation if the literal reading fails. This compulsory literal-first strategy then incurs processing costs for idiomatic or abstract meanings.

Taken together, previous research leaves unresolved a key theoretical question: whether the processing advantage for MWS is primarily a consequence of distributional experience, or whether semantic properties of the expression also shape the processing route that speakers use. Usage-based and statistical learning accounts predict that frequent and familiar sequences become entrenched through repeated exposure, leading to faster processing. On this view, if idioms and collocations are matched on phrase-level frequency, familiarity, and contextual predictability, their processing advantage should be comparable. In contrast, hybrid and dual-route accounts of MWS processing suggest that expressions may differ in the relative contribution of direct retrieval and compositional computation. From this perspective, highly conventionalized expressions, such as idioms, may be more likely to engage direct access to a unitary meaning representation, whereas transparent collocations may rely more strongly on compositional processing, even when they are frequent and familiar.

This contrast is also theoretically important for L2 processing. L2 speakers have less experience with many MWSs and may therefore rely more heavily on literal, word-by-word computation than on direct retrieval of phrase-level meaning. Comparing L1 and L2 speakers therefore allows us to examine whether the balance between direct retrieval and compositional processing differs across language backgrounds. By directly comparing idioms and collocations while controlling distributional factors, the present study tests whether MWS processing is best explained by distributional experience alone, or whether semantic conventionality also shapes how different types of MWSs are represented and processed, and how this varies between L1 and L2 speakers.

Downloaded from https://www.cambridge.org/core. 26 Oct 2026 at 14:05:28, subject to the Cambridge Core terms of use.

The present study

While several studies have demonstrated the processing advantages of MWSs over novel word combinations, the processing differences among different types of MWS remain largely unknown. The present study directly contrasts two main types of MWS, collocations and idioms (used either literally or figuratively) and compares their processing to each other and to novel word combinations. L1 and L2 speakers were tested on the same material. We focused specifically on collocations extracted from corpora and idioms because they represent different points on the scale of semantic conventionality: corpus-defined collocations are typically relatively literal in meaning, whereas idioms are more strongly associated with conventionalised idiomatic meanings that cannot always be derived directly from their component words. Although MWSs include many different expression types and vary widely in syntactic structure, focusing on collocations and idioms allowed us to restrict the comparison to adjective–noun constructions, thereby reducing syntactic variation across item types.

A self-paced reading paradigm was implemented in conjunction with a concurrent stop-making-sense paradigm (Boland et al., 1990; Gibson et al., 1996), to encourage participants to monitor sentence coherence incrementally during reading. This was important because the experimental materials contained collocations and idioms which, despite being selected to be familiar, may still vary in familiarity and interpretability, especially across L2 speakers. The Stop-Making-Sense task allowed us to identify trials in which participants did not treat the sentence as meaningful during unfolding comprehension. Furthermore, the task discourages superficial, "good-enough" processing (Ferreira & Patson, 2007), deriving an acceptable overall interpretation without necessarily fully processing the critical expression.

Our first research question investigates whether semantic conventionality gives idioms a processing advantage over collocations when distributional factors (phrase frequency, familiarity, and contextual predictability) are controlled for. Additionally, it examines whether the processing differences between idioms and collocations vary between L1 and L2 speakers.

We hypothesize that if distributional factors primarily drive the processing advantage of idioms, we should observe no significant differences between idioms and collocations. However, if additional properties of MWS, such as conventionality, play a role in MWS processing, we expect idioms to demonstrate a processing advantage over collocations, particularly among L1 speakers.

The second research question investigates idiom-level variables, i.e., meaning dominance and semantic decomposability, that might influence how L1 and L2 speakers process idioms. As semantic relatedness and decomposability share similarities, we expect an interaction between meaning dominance and semantic decomposability, mirroring Milburn and Warren (2019), with a facilitating effect of semantic decomposability when an idiom is more frequently used in its literal interpretation, whereas an inhibitory effect is expected when an idiom is predominantly used figuratively.

Downloaded from https://www.cambridge.org/core. 26 Oct 2026 at 14:05:28, subject to the Cambridge Core terms of use.

Table 1. Summary of participants' language backgrounds

Measure	L1 speakers *(M, SD)*	L2 speakers *(M, SD)*
Age of onset of English acquisition	1.31 (3.02)	8.05 (3.96)
Time lived in an English-speaking country (in years)	30.5 (6.1)	7.62 (7.84)
Vocabulary Levels Test score (out of 90)	86.2 (4.07)	66.2 (13.4)
Self-reported English proficiency (6-point scale)		
Speaking	4.93 (0.25)	3.59 (0.90)
Listening	4.98 (0.15)	3.89 (0.83)
Reading	4.95 (0.21)	4.00 (0.71)
Writing	4.97 (0.18)	3.62 (0.83)
Self-reported English language frequency of use (hours per day)		
Speaking	8.41 (5.53)	3.57 (3.58)
Listening/Watching	4.24 (3.04)	4.22 (4.69)
Reading	2.60 (2.20)	3.38 (2.57)
Writing	2.93 (2.94)	2.22 (2.06)

Method

Openness and transparency

All study materials, data, and R code required to reproduce the present findings are provided via the Open Science Framework <https://osf.io/fx3mr/>. The present project was not pre-registered.

Participants

Eighty-seven L1 native English speakers (age: $M = 31.90$, $SD = 5.25$) and 117 L1-Chinese speakers of L2-English (age: $M = 29.90$, $SD = 7.62$) participated in the experiment. Four L1 and five L2 participants were removed from the analysis, either for judging nearly all implausible filler sentences as sensible and plausible critical sentences as nonsensical or for scoring below 80% on comprehension questions. Participants' language backgrounds are summarized in Table 1. We conducted a power analysis using simulation-based methods following DeBruine and Barr's (2021) recommendations. Power was evaluated within a null-hypothesis significance testing the faster reading of idioms or collocations over novel word combinations. Effect sizes for the simulations were informed by using collocation studies (Li et al., 2021, 2022) and idiom study (Carrol & Conklin, 2020) as benchmarks for plausible magnitudes. We then generated 1,000 synthetic datasets for a range of candidate sample sizes and, for each dataset, fit a linear mixed-effects model using lmer to log-transformed reading times: lmer(log(dv)) ~ Phrase_type + (Phrase_type | participant) + (1 | item). For each simulation run, we recorded whether the target

Downloaded from https://www.cambridge.org/core. 26 Oct 2026 at 14:05:28, subject to the Cambridge Core terms of use.

fixed effects reached statistical significance. The results indicated that a sample of 80 participants, each reading 40 collocations or idioms and 40 novel word combinations, would provide at least 90% power to detect the critical effects for reading time, using an alpha level of .05.

All participants had normal or corrected-to-normal vision and provided written informed consent. The experiment was conducted in accordance with the British Psychological Society ethical guidelines (version 2021) and was approved by the University of Birmingham's ethics committee (ERN_19-0936).

Additionally, L2 speakers' proficiency was measured using the Vocabulary Levels Test (Schmitt et al., 2001); the average score was 66.2 out of 90. On average, participants had started learning English at 8.05 years ($SD = 3.96$) and had spent 7.62 years in English-speaking countries ($SD = 7.84$). L2 participants completed a multiple-choice test to confirm their knowledge of the target idioms, and any trials with incorrectly interpreted items were excluded from analysis. Previous research normed idiom familiarity on a single L2 cohort and deliberately selected familiar expressions; yet idioms are inherently infrequent (Moon, 1998) and individual exposure varies widely (Kuperman & Van Dyke, 2013). To address this variability, we not only chose idioms reported as familiar but also removed any unknown items based on each participant's post-experiment questionnaire, ensuring that every idiom included in the analyses was genuinely known to our L2 sample.

Design

For the first research question, phrase type was a four-level within-participant and between-item variable (figuratively used idioms, literally used idioms), collocations, and novel word combinations. These three types of phrases reflect varying degrees of semantic conventionality, with idioms exhibiting the most conventionalized meanings, collocations falling in between and novel word combinations lacking a conventionalized meaning. For the second research question, predictors included meaning dominance and local and global decomposability of idioms.

The dependent variables were reading time in the critical and spillover regions. The critical region consisted of the MWS, and the spillover region consisted of the two words that followed the critical region. The spillover region was included because the self-paced reading paradigm can delay processing effects, potentially influencing subsequent regions.

Materials

There were 40 sets of adjective-noun quartets, each sharing the same final word. To select these adjective-noun quartets, we used the following steps. First, we identified 72 adjective-noun pairs listed as idioms in either Macis and Schmitt (2017) or the Oxford Dictionary of English Idioms (Ayto, 2009). Each idiom also needed to be ambiguous, allowing for both idiomatic and literal interpretations. Second, we then extracted 72 adjective–noun collocations from the BNC, each with the target noun immediately preceded by an adjective within a one-word window. All collocations had a Mutual Information (MI) score of 3 or higher (Hunston, 2022). Third, we

Downloaded from https://www.cambridge.org/core. 26 Oct 2026 at 14:05:28, subject to the Cambridge Core terms of use.

generated novel adjective–noun pairs by pairing each target noun with a literal adjective attested at most once or never in the BNC. Finally, to norm these items, 45 L1-English speakers rated the 72 sets on familiarity using a scale from 1 (Never) to 5 (Extremely frequently), indicating how often they encountered or used each item. Based on these ratings, 40 sets of items were selected. Table 2 reports the item characteristics.

Item-related variables were matched as closely as possible across the four items in each set. One-way repeated ANOVA tests indicated that adjective frequency was matched across conditions ($F(2, 117) = 1.22$, $p = .30$). However, adjective length differed, resulting in differences in phrase-level length ($F(2, 117) = 3.55$, $p = .032$), with idioms being shorter than novel word combinations ($\beta = 0.7$, $t(117) = 2.46$, $CI = [0.14, 1.26]$, $p = .046$). Note that the difference was less than 1 character on average. Finally, idioms and collocations were matched on phrase-level frequency ($t(39) < 1$, $CI = [-28.71, 22.66]$), MI-score ($t(39) < 1$, $CI = [-1.26, 1.11]$), and familiarity ($t(39) < 1$, $CI = [-0.51, 0.18]$).

Idioms: meaning dominance and decomposability

The meaning dominance and local and global decomposability of idioms were assessed by the same 45 L1-English speakers who provided familiarity ratings. First, they evaluated figurative meaning dominance on a slider, indicating the percentage of times they anticipated the idiom would be understood literally or figuratively, with 0% meaning exclusively literal and 100% meaning exclusively figurative (see Milburn & Warren, 2019). These ratings demonstrated high reliability (Cronbach's alpha = 0.912). Their consistency was evidenced by high correlations with estimates from another 20 L1-English speakers ($r(38) = 0.7$, $CI = [0.52, 0.83]$, $p < .001$).

Then, the same 45 speakers assessed *local decomposability*, by rating the contributions of the adjective and the noun to the idiom's figurative meaning on a scale from 1 (no contribution) to 7 (100% contribution; see Gibbs et al., 1989). *To assess global decomposability*, an additional 15 L1-English speakers rated the literal meaning of the adjective-noun pair to contribute to the idiom's figurative meaning, using the same scale. These measures are reported in Table 2.

Context sentences

For each adjective-noun quartet, four context sentences were created. Each context sentence consisted of a disambiguating context, a control region, a critical region containing the adjective-noun combination, a spillover region (the two words following the critical region), and a wrap-up region, as shown in Table 3. The control and spillover regions were identical across conditions for each set of items. For idioms, disambiguating contexts were biased toward either the literal or figurative interpretation.

To measure contextual predictability, 60 L1 English speakers completed a sentence continuation task, predicting the noun of 40 sentences truncated before the noun of the MWS (counterbalanced over four lists, with an equal number of items per condition). Contextual predictability for the final nouns was matched among

Downloaded from https://www.cambridge.org/core. 26 Oct 2026 at 14:05:28, subject to the Cambridge Core terms of use.

Downloaded from https://www.cambridge.org/core. 26 Oct 2026 at 14:05:28, subject to the Cambridge Core terms of use.

Table 2. Summary of item characteristics (M, SD)

Items characteristics	Novel word combinations	Collocations	Idioms-overall	Idioms-literal	Idioms-figurative
Adjective length	5.1 (1.06)	5 (1.04)	4.42 (0.90)	–	–
Phrase-level length	9.55 (1.24)	9.45 (1.24)	8.85 (1.33)	–	–
Adjective frequency	10825 (8789)	13260 (13392)	15186 (14607)	–	–
Phrase-level frequency	–	51.7 (51.4)	54.7 (55.4)	–	–
Familiarity (1–5)	–	3.29 (0.8)	3.45 (0.66)	–	–
Predictability	0.11 (0.17)	0.08 (0.12)	–	0.07 (0.11)	0.04 (0.12)
Sentence naturalness	3.9 (0.66)	4.06 (0.5)	–	4.06 (0.52)	4.18 (0.61)
Decomposability: adjectives (1–7)	–	–	5.21 (0.79)	–	–
Decomposability: nouns (1–7)	–	–	4.1 (0.89)	–	–
Overall decomposability (1–7)	–	–	4.48 (0.63)	–	–
Figurative dominance (0–100%)	–	–	69% (16.3)	–	–

Note: All frequencies were extracted from the 100-million-word BNC and are presented as raw occurrence counts.

Table 3. An example of one set of items and context sentences

	Sentences				
Condition	Disambiguating region	Control region	Critical region	Spillover region	Wrap-up region
novel word combination	Although most of it is known to us,		Rare area		
collocation	As Russia is such a big country,	there is a	Vast area	that we	have not noticed.
idiom-literal	As the wall is not painted properly,		Gray area		
idiom-figurative	As the law is not totally transparent,		Gray area		

literal idioms, figurative idioms, and collocations ($F(2, 117) = 1.2$, $p = .3$). To ensure sentence readability, reading times up to the target regions were compared across the three conditions, and no differences were found ($\chi 2(1) = 3.33$, $p = .35$). Second, an additional 20 L1 English speakers rated the naturalness of the sentence frames (up to the spillover regions) on a five-point scale (1 = *very unnatural*, 5 = *very natural*), and no differences were found across conditions (F (3, 156) = 1.67, $p = .18$). To ensure that participants monitored sentence meaning in the stop-making-sense task, 40 semantically incorrect or odd filler sentences were included. These varied in implausibility location and were judged implausible by over 80% of 20 L1-English speakers, who normed them alongside 20 grammatically correct and semantically plausible sentences.

After all pre-tests, four lists of 80 trials were created by counterbalancing the 40 critical sets of items following a Latin square design and including the 40 filler sentences. The software randomized each list for each participant and presented the trials in two 40-trial blocks separated by a self-paced break.

Procedure

Participants were recruited via Prolific and Sona Systems and directed to the web-based reading experiment. The experiment was programmed with PsychoPy (version 2021.2.3) and hosted by Pavlovia. Participants read instructions on the computer screen and then completed eight practice trials. At the start of each trial, the prompt "A new sentence starts" was displayed and disappeared once the participant pressed the space bar. Each sentence was divided into five reading regions, with the critical region always appearing third. Reading regions were presented one at a time in the centre of the screen to ensure the same position across participants' different screen sizes. Participants initiated reading by pressing the space bar to reveal the first region. To progress through the sentence, they pressed "J" if the text remained plausible, which removed the current region and replaced it with the next region at the same central location. If the sentence became nonsensical for them, they pressed "F," which terminated the trial and initiated the next one. If the participant read the full sentence and rated it as plausible, a comprehension question was presented to assess

Downloaded from https://www.cambridge.org/core. 26 Oct 2026 at 14:05:28, subject to the Cambridge Core terms of use.

understanding. After the reading experiment, all participants completed a language background questionnaire.

Statistical analysis

All analyses were carried out using R version 4.2.1 (R Core Team, 2022) with *lm4* (Bates et al., 2015). Following recommendations (Brysbaert & Stevens, 2018) for analyzing continuous data, RTs were log-transformed prior to analysis to better approximate a normal distribution. For the first question, the full model included phrase type and language background as fixed effects, and trial number, phrase length, adjective, and noun frequencies as covariates. We treatment-coded phrase type with novel word combinations as 0 and language background with L1 as 0. For the second question, the full model included language background, idiom-level variables (meaning dominance, global decomposability, adjective decomposability, and noun decomposability), and the interactions between idiom-level variables and the same covariates as the first research question. All continuous predictors were cantered before entering them into models. Finally, to test whether L2 proficiency modulates phrase type effect, a separate model was fit to L2 speaker data only, including vocabulary scores, phrase type, and their interaction with the same covariates as the first model.

In all models, the maximal random structure model included the interactions of experimental variables with both participants and items whenever possible (Barr et al., 2013). If convergence was not achieved, the "bobyqa" or "optimx" optimiser was used to increase the number of iterations. If this failed, the random effect was simplified by removing interactions first, followed by individual terms if necessary, until the model converged.

All full models were subsequently reduced with the drop1() function, which compares the model with a chi-square test. The manipulated variables were always retained in the corresponding reduced models to reflect the experimental design, even if their effects were not significant; covariates were removed if not significant. Based on these tests, reduced models were selected to best fit the data. For all bested-fitted models, the vif() function from the *car* package was used to check collinearity and variables with a t value of greater than 1.96 were considered significant (Baayen et al., 2008). Finally, when there was a significant interaction, the emmeans function from the emmeans package (Lenth et al., 2024) was used to conduct pairwise comparisons of estimated marginal means using Tukey-adjusted p-values.

Data pre-processing

Following Singh et al. (2016), critical trials that participants thought implausible in English were removed, resulting in 15% L1 and 20% L2 data loss. To ensure that the data removal did not affect the results, we also ran the models with the excluded data points, showing similar results in terms of direction and significance. Trials associated with incorrect responses to comprehension questions (3% of L1 data and 5% of L2 data) were excluded from the analysis. An extra 1.8% of L2 data were removed because the participant failed to identify the idiom's meaning in the multiple-choice

Downloaded from https://www.cambridge.org/core. 26 Oct 2026 at 14:05:28, subject to the Cambridge Core terms of use.

test. RTs above 3000 ms and below 200 ms in the critical region, and those above 2000 ms and below 200 ms in the spillover region were removed (as the spillover region was shorter, the maximum cut-off was lower). Finally, RTs more than 2.5 *SD*s above or below the mean for each participant and condition were excluded. After data cleaning, 80% of the critical region data (2756 trials) and 80% of the spillover region data (2774 trials) from L1 speakers; 72% of the critical region data (3347 trials), and 72% of the spillover region data (3318 trials) from L2 speakers were included in the analyses.

Results

The average reading times in the critical and spillover regions are reported in Table 4. Tables 5 and 6 present the best-fitting models for the different types of MWS and the idiom-only analyses. L2 results are reported in Table 7.

Research question 1.

Our first research question asks whether semantic conventionality gives idioms a processing advantage over collocations after controlling for distributional factors and whether this effect differs between L1 and L2 speakers. In the critical region, we did not observe an interaction between language background and phrase type ($\chi 2(3) = 3.11$, $p = 0.374$). As expected, there was an effect of language background, with L2 speakers showing slower RTs than L1 speakers ($\beta = 0.15$, $t = 3.77$, CI = [0.073, 0.233], $p < .001$). We also found an effect of phrase type. Specifically, collocations were processed faster than novel word combinations ($\beta = -0.107$, $t = -3.4$, CI = [−0.169, −0.045], $p < .001$), and idioms in a figurative context were processed faster than novel word combinations ($\beta = -0.099$, $t = -3.1$, CI = [−0.162, −0.036], $p = .002$).

In the spillover region, there was a significant interaction between language background and phrase type ($\chi 2(3) = 6.65$, $p = 0.036$). Phrase type remained significant for L1 speakers but not for L2 speakers (see Figure 1). Further, pairwise comparisons revealed that L1 speakers processed figuratively used idioms faster than collocations ($\beta = -0.07$, $t = -2.97$, $p = .008$), but L2 speakers processed figuratively used idioms more slowly than literally used idioms ($\beta = 0.03$, $t = 2.81$, $p = .02$).

Finally, to examine the effect of L2 proficiency, we fit a model to the subset of L2 speakers only. In the critical region, L2 speakers exhibited a trend toward faster reading of idioms when presented in figurative contexts ($\beta = -0.06$, $t = -1.921$, $p = .057$) and demonstrated significantly faster reading times for collocations compared to novel word combinations ($\beta = -0.08$, $t = -2.51$, $p = .012$). Vocabulary score reached significance, with L2 speakers with higher vocabulary scores showing faster reading times ($\beta = -0.004$, $t = -1.98$, $p = .048$), but no significant interaction was found between vocabulary score and phrase type ($p = 0.296$). In the spillover region, an interaction between phrase type and vocabulary score was observed ($\chi 2(3) = 8.35$, $p = .039$). To further explore this interaction, we examined the simple effect of phrase type at high and low vocabulary

Downloaded from https://www.cambridge.org/core. 26 Oct 2026 at 14:05:28, subject to the Cambridge Core terms of use.

Table 4. Mean RTs (ms) across four conditions by language background

Language background	Condition	Critical region (*M, SD*)	Spill-over region (*M, SD*)
L1 speaker	Novel word combination	1178 (510)	764 (284)
	Collocation	1035 (422)	782 (300)
	Idiom in literal use	1096 (473)	745 (273)
	Idiom in figurative use	1062 (473)	738 (288)
L2 speaker	Novel word combination	1394 (634)	772 (269)
	Collocation	1272 (584)	777 (296)
	Idiom in literal use	1307 (594)	731 (253)
	Idiom in figurative use	1274 (600)	769 (305)

levels (See Figure 2). The lower-vocabulary group showed significantly longer RTs for figurative than literal idioms ($\beta = 0.05$, $p = .013$).

Research question 2.

The second research question investigates how idiom-level properties influence L1 and L2 idiom processing. Only the semantic decomposability of the noun had an effect ($\beta = -0.07$, $t = -2.42$, $p = .02$), with higher semantic decomposability speeding up the reading of idioms in the critical region.

Finally, we fitted a separate model to the L2 speaker data. Noun semantic decomposability and vocabulary knowledge were the only predictors of idiom RTs, and only in the critical region, where higher noun decomposability ($\beta = -0.07$, $t = -2.48$, $p = .013$) and higher vocabulary scores ($\beta = -0.005$, $t = -2.26$, $p = .024$) were associated with shorter idiom RTs.

General discussion

Previous studies have demonstrated faster processing of idioms and collocations among L1 speakers and faster processing of collocations among L2 speakers. However, differences in the processing of various types of MWSs remain unexplored. The present study aimed to compare the processing of different types of MWSs in L1 and L2 speakers. First, we replicated previous findings that both L1 and L2 speakers read collocations more quickly than novel word combinations. Similarly, L1 speakers processed idioms used figuratively faster than novel word combinations. Importantly, our study revealed that, on the spillover region, only L1 speakers continued to process figurative idioms more quickly than collocations, even when controlling for distributional factors such as phrase-level frequency, familiarity, and contextual predictability. In contrast to L1 speakers, who demonstrated faster processing of figurative idioms compared to literal idioms, lower proficiency L2 speakers exhibited faster processing of literal idioms compared to figurative idioms.

Downloaded from https://www.cambridge.org/core. 26 Oct 2026 at 14:05:28, subject to the Cambridge Core terms of use.

Downloaded from https://www.cambridge.org/core. 26 Oct 2026 at 14:05:28, subject to the Cambridge Core terms of use.

Table 5. Summary of the best-fitting models of MWSs on critical and spill-over regions

Critical region						Spillover region					
Predictors	β	SE	CI	t	p	Predictors	β	SE	CI	t	p
(Intercept)	7.00	0.035	6.93–7.07	200.25	<.001	(Intercept)	6.61	0.027	6.55–6.66	241.05	<.001
Trial number	−0.002	0.000	−0.003–−0.002	−7.47	<.001	Trial number	−0.003	0.000	−0.003–−0.002	−11.61	<.001
Item length	0.013	0.009	−0.005–0.031	1.42	.159	Phrase type [idiom figurative]	−0.056	0.024	−0.104–−0.009	−2.33	.021
Noun frequency	−0.045	0.012	−0.068–−0.022	−3.90	<.001	Phrase type [idiom literal]	−0.029	0.024	−0.077–0.019	−1.21	.230
Adjective frequency	−0.027	0.012	−0.050–−0.004	−2.28	.024	Phrase type [collocation]	0.018	0.024	−0.029–0.065	0.77	.443
Phrase type [idiom figurative]	−0.099	0.032	−0.162–−0.036	−3.10	.002	Language background [L2]	0.012	0.033	−0.052–0.076	0.37	.709
Phrase type [idiom literal]	−0.064	0.032	−0.127–−0.001	−2.00	.047	Item length	−0.010	0.007	−0.024–0.005	−1.33	.187
Phrase type [collocation]	−0.107	0.031	−0.169–−0.045	−3.40	<.001	Phrase type [collocation] × Language background [L2]	−0.014	0.022	−0.058–0.030	−0.63	.530
Language background [L2]	0.153	0.041	0.073–0.233	3.77	<.001	Phrase type [idiom figurative] × Language background [L2]	0.047	0.022	0.003–0.091	2.09	.038
Phrase type [idiom figurative] × Language background [L2]	0.046	0.029	−0.011–0.102	1.60	.112	Phrase type [idiom literal] × Language background [L2]	−0.013	0.022	−0.057–0.030	−0.61	.543

(*Continued*)

Downloaded from https://www.cambridge.org/core. 26 Oct 2026 at 14:05:28, subject to the Cambridge Core terms of use.

Table 5. (*Continued*)

Critical region						Spillover region					
Predictors	β	SE	CI	t	p	Predictors	β	SE	CI	t	p
Phrase type [idiom literal] × Language background [L2]	0.019	0.028	−0.036–0.074	0.67	.502						
Phrase type [collocation] × Language background [L2]	0.031	0.028	−0.024–0.087	1.11	.269						
Random effects						**Random effects**					
σ^2					0.10	σ^2					0.05
τ_{00} ID					0.06	τ_{00} ID					0.04
τ_{00} items					0.01	τ_{00} items					0.01
ICC					0.44	ICC					0.47
N_ID					204	N_ID					203
τ_{11} items.Language_backgroundL2					0.01	τ_{11} items.Language_backgroundL2					0.00
ρ_{01} items					−0.26	ρ_{01} items					0.15

Downloaded from https://www.cambridge.org/core. 26 Oct 2026 at 14:05:28, subject to the Cambridge Core terms of use.

Table 6. Summary of the best-fitting models of idioms on critical and spill-over regions

Critical region						Spillover region					
Predictors	β	SE	CI	t	p	Predictors	β	SE	CI	t	p
(Intercept)	7.364	0.194	6.984–7.744	37.98	<.001	(Intercept)	6.641	0.139	6.368–6.915	47.65	<.001
Adjective frequency	−0.049	0.02	−0.088–−0.010	−2.43	.020	Meaning dominance	0.04	0.028	−0.015–0.095	1.42	.163
Noun frequency	−0.054	0.024	−0.100–−0.007	−2.27	.029	Noun decomposition	−0.034	0.019	−0.071–0.003	−1.79	.080
Noun decomposition	−0.07	0.029	−0.128–−0.013	−2.42	.020	Language background (L2)	0.092	0.068	−0.042–0.226	1.35	.177
Language background (L2)	0.084	0.087	−0.086–0.255	0.97	.333	Adjective decomposition	0.024	0.023	−0.022–0.070	1.01	.320
Adjective decomposition	−0.022	0.034	−0.089–0.045	−0.64	.528	Global decomposition	−0.024	0.029	−0.082–0.034	−0.82	.420
Global decomposition	−0.019	0.041	−0.100–0.062	−0.46	.645	Trial number	−0.003	0.001	−0.004–−0.002	−5.5	<.001
Meaning dominance	0.024	0.041	−0.056–0.105	0.59	.556	Noun decomposition ×Language background (L2)	−0.01	0.014	−0.038–0.018	−0.71	.479
Trial number	−0.003	0.001	−0.005–−0.002	−4.1	<.001						
Noun decomposition × Language background (L2)	0.025	0.019	−0.011–0.061	1.35	.179						
Random effects						**Random effects**					
σ^2	0.1					σ^2	0.05				
τ_{00} ID	0.06					τ_{00} ID	0.04				
τ_{00} items	0.01					τ_{00} items	0.01				
ICC	0.44					ICC	0.47				
N_ID	204					N_ID	203				
τ_{11} items Language_backgroundL2	0.01					τ_{11} items Language_backgroundL2	0				
ρ_{01} items	−0.26					ρ_{01} items	0.15				

Downloaded from https://www.cambridge.org/core. 26 Oct 2026 at 14:05:28, subject to the Cambridge Core terms of use.

Table 7. The separate best-fitting models for L2 speakers

MWS for L2 speakers						Idioms for L2 speakers					
	Critical region						Critical region				
Predictors	β	SE	CI	t	p	Predictors	β	SE	CI	t	p
(Intercept)	7.16	0.03	7.1–7.23	214.85	<0.001	(Intercept)	7.44	0.21	7.02–7.86	34.78	<0.001
Phrase type [figurative idiom]	−0.06	0.03	−0.12–0.001	−1.92	0.055	Adjective frequency	−0.05	0.02	−0.10–−0.01	−2.41	0.016
Phrase type [collocation]	−0.08	0.03	−0.14–−0.02	−2.51	0.012	Adjective decomposition	0.01	0.04	−0.06–0.09	0.34	0.738
Phrase type [literal idiom]	−0.056	0.03	−0.12–−0.006	−1.78	0.074	Noun decomposition	−0.07	0.03	−0.12–−0.01	−2.48	0.013
Trial number	−0.002	0.0004	−0.003–−0.001	−4.93	<0.001	Global decomposition	−0.04	0.05	−0.13–0.05	−0.91	0.363
Noun frequency	−0.049	0.013	−0.057–−0.007	−3.84	<0.001	Meaning dominance	0.05	0.05	−0.04–0.14	1.09	0.278
Adjective frequency	−0.032	0.013	−0.057–−0.007	−2.51	0.012	Vocabulary	−0.005	0.002	−0.01–−0.001	−2.26	0.024
Vocabulary	−0.004	0.002	−0.007–0	−1.98	0.048	Trial number	−0.002	0.001	−0.004–0	−2.03	0.042
Random effects						**Random effects**					
σ^2	0.1					σ^2	0.09				
τ_{00} items	0.02					τ_{00} ID	0.08				
τ_{00} ID	0.07					τ_{00} items	0.01				
ICC	0.45					ICC	0.49				
N_ID	117					N_ID	115				
N_items	120					N_items	39				
Observations	3256					Observations	708				
Marginal R^2/ Conditional R^2	0.047/ 0.478					Marginal R^2/ Conditional R^2	0.074/ 0.528				

(*Continued*)

Downloaded from https://www.cambridge.org/core. 26 Oct 2026 at 14:05:28, subject to the Cambridge Core terms of use.

Table 7. (*Continued*)

MWS for L2 speakers						Idioms for L2 speakers					
Spillover region						Spillover region					
Predictors	β	SE	CI	t	p	Predictors	β	SE	CI	t	p
(Intercept)	6.62	0.03	6.56–6.67	250.53	<0.001	(Intercept)	6.77	0.17	6.43–7.11	39.06	<0.001
Phrase type [collocation]	0.002	0.03	−0.05–0.06	0.07	0.945	Adjective decomposition	0.03	0.03	−0.02–0.09	1.17	0.241
Phrase type [literal idiom]	−0.036	0.03	−0.1–0.02	−1.36	0.175	Noun decomposition	−0.04	0.02	−0.08–−0.002	−1.88	0.061
Phrase type [figurative idiom]	−0.003	0.03	−0.06–0.06	−0.093	0.926	Vocabulary	-0.002	0.001	−0.01–0.001	−1.46	0.145
Vocabulary	−0.002	0.001	−0.005–−0.001	−1.52	0.128	Meaning dominance	0.05	0.04	−0.02–0.12	1.49	0.138
Trial number	−0.003	0.0003	−0.003–−0.002	−7.61	<0.001	Global decomposition	−0.05	0.04	−0.12–0.02	−1.35	0.178
Random effects						Trial number	−0.003	0.001	−0.004–−0.001	−3.87	<0.001
σ^2	0.06					**Random effects**					
τ_{00} items	0.01					σ^2	0.06				
τ_{00} ID	0.03					τ_{00} ID	0.04				
ICC	0.44					τ_{00} items	0.01				
N_ID	116					ICC	0.43				
N_items	120					N_ID	114				
Observations	3318					N_items	39				
Marginal R^2/ Conditional R^2	0.040/ 0.467					Observations	731				
						Marginal R^2/ Conditional R^2	0.082/ 0.474				

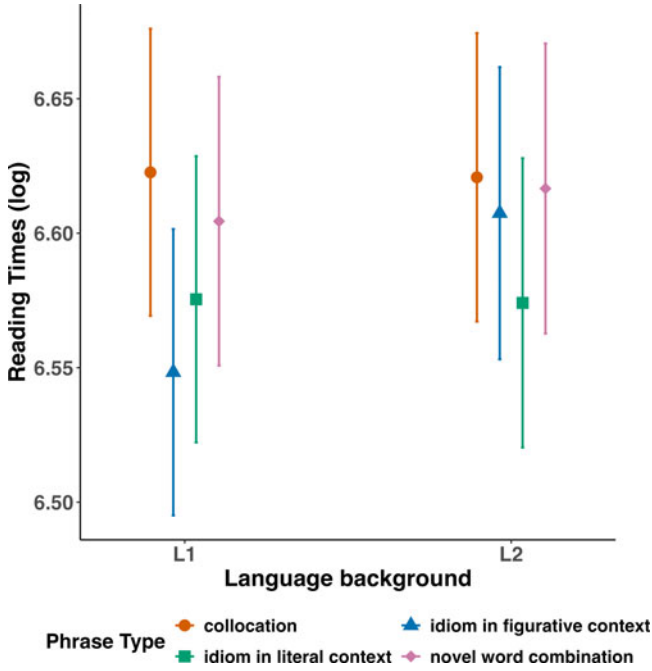


Figure 1. The interaction between language background and phrase type.

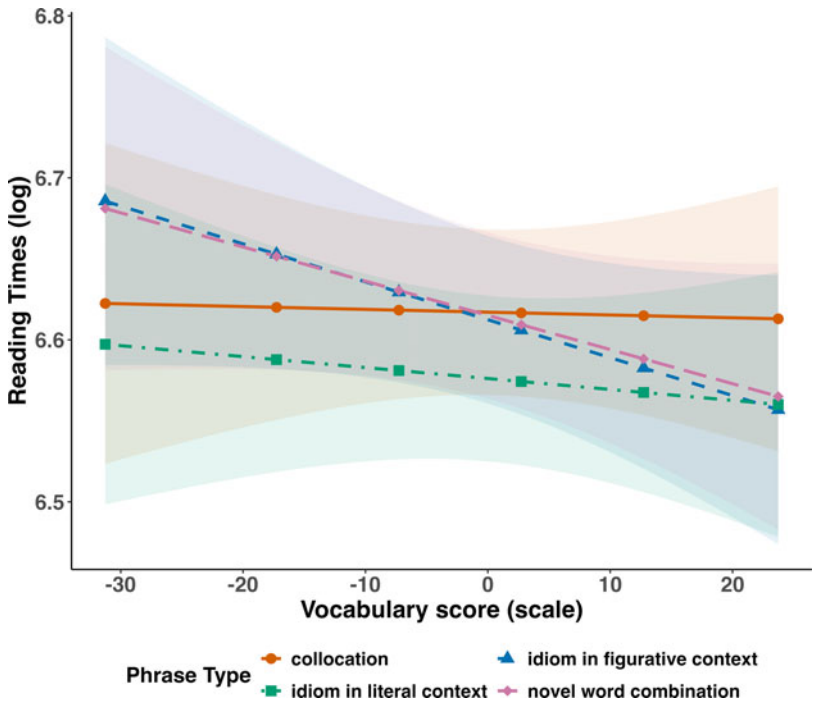


Figure 2. The interaction between phrase type and vocabulary score.

Downloaded from https://www.cambridge.org/core. 26 Oct 2026 at 14:05:28, subject to the Cambridge Core terms of use.

The distributional information effect

Using self-paced reading in conjunction with a concurrent stop-making-sense paradigm (Boland et al., 1990; Gibson et al., 1996), we replicated findings from other paradigms showing faster processing of collocations compared to novel word combinations, in both L1 and L2 speakers (e.g., Li et al., 2025; Li et al., 2021, 2022; Öksüz et al., 2021; Sonbul, 2015; Vilkaitė, 2016; Vilkaitė & Schmitt, 2019). Our study further demonstrated that L1 speakers read figuratively used idioms faster than novel word combinations, consistent with previous findings in L1 idiom processing (e.g., Carrol et al., 2016; Carrol & Conklin, 2017; Siyanova-Chanturia et al., 2011; Swinney & Cutler, 1979; Tabossi et al., 2009). These results suggest that language users exhibit a processing advantage for collocations and idioms, indicating that these word combinations become entrenched in memory over time. Consequently, this advantage primarily arises from the higher phrase-level frequency and greater familiarity of collocations and idioms (Carrol & Conklin, 2020), extending the frequency effect beyond single words to multiword sequences. Moreover, these findings align with language models that emphasize the role of experience—specifically, the notion that frequency serves as an index of speakers' experience with a linguistic construction, which impacts language entrenchment and processing. This perspective is supported by usage-based models of language acquisition (Bybee, 1998; Goldberg, 2006; Tomasello, 2003), which argue that linguistic knowledge is shaped by repeated exposure and experience. The processing advantage in collocations and idioms over novel word combinations also supports statistical learning theories, which posit that this advantage reflects language users' sensitivity to word co-occurrence. This is consistent with the view that both L1 and L2 speakers are capable of tracking patterns of word co-occurrence, as demonstrated in the statistical learning of language (Rebuschat & Williams, 2012; Romberg & Saffran, 2010; Saffran & Kirkham, 2018; Wang & Saffran, 2014).

Beyond distributional information

We first examined whether idioms exhibit a processing advantage over collocations once distributional information (i.e., phrase-level frequency, familiarity, and contextual predictability) is controlled for. The spill-over region revealed that L1 speakers read figuratively used idioms faster than collocations, suggesting that the processing advantage of idioms extends beyond distributional information. If MWSs such as collocations and idioms become familiar and predictable through repeated exposure (e.g., Cacciari & Tabossi, 1988; Libben & Titone, 2008; Tabossi et al., 2009) and if distributional information alone drives the processing advantage, then collocations and idioms should not differ in our experiment, given that phrase-level frequency, familiarity, and contextual predictability were controlled for. Hence, the idiom advantage in the spill-over region suggests a processing effect beyond distributional information.

One reason for the advantage of figuratively used idioms over collocations could be that idioms tend to have more conventionalised meanings. Because idiomatic meanings cannot simply be computed by assembling the meanings of their lexical components, idioms are generally less decomposable and more conventionalised. In

Downloaded from https://www.cambridge.org/core. 26 Oct 2026 at 14:05:28, subject to the Cambridge Core terms of use.

contrast, collocations are more decomposable and less conventionalised, as their collocational meaning can be computed based on their lexical components (all collocations used in our experiments were semantically transparent and literal).

According to the Hybrid Representation Hypothesis (Kuiper et al., 2007; Sprenger et al., 2006), MWSs, e.g., idioms, can be activated both through the analysis of their lexical parts and via direct access to the unitary meaning, and this dual activation can occur interactively and synchronously. Our findings extend this hypothesis by suggesting that the relative contribution of these two routes is not fixed but graded along a continuum determined by the degree of conventionality. Specifically, highly conventionalised MWSs such as figurative idioms are more strongly represented as holistic units and therefore more likely to be processed via direct retrieval of their conventional meaning. Collocations, while frequent, are more semantically transparent and thus more likely to rely on compositional analysis, even though direct access may also be available. Thus, rather than assuming categorical processing differences between idioms and collocations, our results support a graded relationship between the direct retrieval route and the computational route.

We therefore propose that the representation of MWSs and their internal components is redundant, with both holistic and decomposed information co-existing but differing in cognitive prominence. As the degree of conventionality increases, holistic representations become more dominant, and processing shifts toward direct retrieval. As conventionality decreases, processing relies more heavily on sequential compositional computation. This graded account integrates the Hybrid Representation Hypothesis within a broader continuum framework and provides an explanation for why idioms, but not collocations, retain a processing advantage after distributional factors are controlled.

Processing idioms in a second language

Unlike some previous studies on L2 idiom processing (e.g., Carrol et al., 2016; Carrol & Conklin, 2017; Siyanova-Chanturia et al., 2011), this study found that L2 speakers exhibited a trend toward faster recognition of figurative idioms compared to novel word combinations. The discrepancy between prior research and the current study may stem from the exclusion of unfamiliar idioms for L2 speakers. While we selected idioms that were rated familiar in prior research, idioms are relatively infrequent in language (Moon, 1998) and individual variability is high for less common linguistic constructions (Kuperman & Van Dyke, 2013). We therefore excluded idioms that participants reported as unknown in a post-experiment task from analyses. Hence, the L2 reading data were not confounded by participants not knowing idiomatic interpretations.

According to the lexical entrenchment hypothesis (Diependaele et al., 2013), greater exposure to linguistic structures in the L2 strengthens both representations and the links between them, reducing activation thresholds and facilitating processing of frequent idioms. However, idiom representations in L2 speakers, while entrenched, may remain weaker than those in L1 speakers, especially among speakers with lower vocabulary levels. The interaction between phrase type and

Downloaded from https://www.cambridge.org/core. 26 Oct 2026 at 14:05:28, subject to the Cambridge Core terms of use.

vocabulary level in the spill-over region indicated a processing advantage for idioms used in a literal context by weaker L2 speakers but not by more proficient L2 speakers. In contrast, L1 speakers showed a processing advantage for figuratively used idioms.

This asymmetry is theoretically informative. Within the Hybrid Representation Hypothesis (Kuiper et al., 2007; Sprenger et al., 2006), both direct retrieval and compositional analysis are available for idiom processing. Our findings suggest that L2 speakers operate within the same dual-route architecture as L1 speakers, but the relative weighting of the two routes differs. Specifically, because holistic idiom representations are less strongly entrenched in L2, the direct retrieval route is less dominant, and compositional analysis plays a more central role.

This interpretation is consistent with previous findings showing stronger competition between idiomatic and literal meanings in L2 processing (Carrol & Conklin, 2017; Shi et al., 2023; Siyanova-Chanturia et al., 2011). Furthermore, unlike L1 speakers, L2 speakers did not read figuratively used idioms faster than collocations. Combined with the relative ease of processing literally used idioms, the accumulated evidence indicates that L2 speakers do not strongly activate figurative meanings, even when supported by a biasing figurative context. Instead, L2 speakers appear to prioritize computation of literal meanings, although this tendency becomes less pronounced for more advanced L2 speakers. Indeed, the pattern exhibited by more advanced L2 speakers in our study began to trend toward the L1 pattern, though without fully converging.

Rather than treating this difference as evidence for qualitatively distinct processing mechanisms in L1 and L2, we argue that it reflects a dynamic shift along the graded retrieval–computation continuum proposed above. In L1 speakers, extensive exposure and early acquisition result in strongly entrenched holistic idiom representations. As a result, the direct retrieval route tends to dominate, particularly for highly conventionalized figurative idioms. In contrast, in L2 speakers, especially those with lower proficiency, the weaker entrenchment of idiomatic representations shifts processing toward greater reliance on compositional computation. As exposure and proficiency increase, the balance gradually shifts toward stronger holistic activation, bringing L2 processing closer to the L1 end of the continuum.

This account integrates our findings with the Literal Salience Model (Cieślicka, 2006), which proposes that literal meanings are more readily activated in L2 processing. Under our graded framework, literal salience reflects the relative strength of compositional routes when holistic idiom representations are insufficiently entrenched. Thus, literal-first tendencies in L2 are not obligatory or categorical, but probabilistic and experience-dependent. In general, when it comes to known idioms, L1 speakers primarily rely on direct retrieval, with computational analysis playing a less important role. In contrast, compositional analysis serves as the primary route for idiom processing in L2 speakers, while direct retrieval functions as a secondary route. Two key reasons may explain this disparity between L1 and L2 speakers. First, L2 speakers have comparatively less exposure and experience with L2 idioms, especially since individual idioms are relatively uncommon in language. This limited exposure makes it more challenging for them to establish strong form – meaning connections for idioms, thereby reducing their ability to identify idioms immediately (Carrol & Conklin, 2017; Senaldi & Titone,

Downloaded from https://www.cambridge.org/core. 26 Oct 2026 at 14:05:28, subject to the Cambridge Core terms of use.

2024). Second, L2 speakers generally follow a more compositional approach than L1 speakers (Beck & Weber, 2016; Cieślicka, 2006; Senaldi & Titone, 2022). Supporting this view, rating tasks (Hubers et al., 2020) show that L2 speakers perceive idioms as more decomposable and transparent than L1 speakers. Consequently, L2 speakers must exert additional effort to inhibit inappropriate literal meanings due to their stronger reliance on compositional analysis.

Crucially, our findings suggest that these L2 characteristics can be understood as reflecting a shift in the relative weighting of processing routes within a shared architecture. Degree of semantic conventionality determines the structural position of an expression along the continuum, while language experience determines where the speaker operates on that continuum. Figurative idioms in L1 speakers occupy the end characterised by dominant direct retrieval; figurative idioms in lower proficiency L2 speakers shift toward greater reliance on computation; and increasing proficiency gradually moves L2 speakers toward the L1-like balance.

Effects of idiom variables

Idiom-level variables, such as meaning dominance and semantic decomposability, had a limited impact on idiom processing. No effect of meaning dominance was observed in this study, indicating that idioms with stronger figurative dominance were not processed faster than those with more balanced meanings, unlike the findings of Milburn and Warren (2019). This discrepancy may be due to the narrower range of meaning dominance in our idioms (63.6–73.6) compared to their study (50.17–89.64). The semantic decomposability of the nouns was found to facilitate idiom processing. This suggests that when the idiomatic and literal meanings of the nouns are more closely related, processing is generally easier for both L1 and L2 speakers, aligning with some prior research (Caillies & Butcher, 2007; Carrol & Segaert, 2024; Gibbs & Nayak, 1989; Hamblin & Gibbs, 1999; Libben & Titone, 2008; Milburn et al., 2021; Titone et al., 2019). This effect of semantic decomposability likely reflects the role of relational overlap between the literal meanings of the constituents and their idiomatic meaning. When such overlap exists, both L1 and L2 speakers can draw on compositional mechanisms to support idiom processing, in addition to retrieving the idiomatic meaning.

Limitations and future research

There are inevitable limitations to our study. First, although the experiment was designed to control many item-related variables, the length of the idioms was somewhat short than that of novel word combinations (though less than 1 character). To address this, the models included item length as a covariate. Nevertheless, this issue should be more rigorously controlled in future studies. Another limitation involved the online experimental paradigm. While online experiments offer advantages such as greater participant diversity and larger sample sizes, data quality remains a concern. To address this, we used a stop-making-sense task, which is sensitive to sentence processing (Boland et al., 1990; Gibson et al., 1996; Singh et al.,

Downloaded from https://www.cambridge.org/core. 26 Oct 2026 at 14:05:28, subject to the Cambridge Core terms of use.

2016). This approach allowed participants to stop reading when they faced comprehension difficulties, particularly with idioms. However, it led to a higher rate of trial exclusions than standard self-paced reading tasks. We justified excluding these trials for two reasons: pressing "NO" indicated comprehension difficulties, especially for L2 speakers, and switching from "Yes" to "No" required extra time, making their inclusion problematic. Importantly, additional analyses including these excluded trials showed similar results, confirming the robustness of our findings. Finally, the self-paced reading paradigm approximates rather than fully replicates natural reading, as sentences are presented in discrete regions rather than as continuous text. Future research using eye-tracking would allow participants to read more naturally while still providing fine-grained measures of MWS processing.

Both L1 and L2 speakers processed collocations more quickly than literal controls. However, while L1 speakers demonstrated a processing advantage for idioms over collocations, this advantage was not observed in L2 speakers. This suggests that frequency alone may not fully account for the learning and processing difficulties in MWS in L2 speakers. Future research should, therefore, consider the semantic aspects of MWS as well. Previous studies have shown that words and even collocations can be gradually and incidentally learned through repeated encounters in context (Chen & Truscott, 2010; Webb, 2007; Webb et al., 2013). However, our findings suggest that repeated exposure to idioms may not be sufficient for effective L2 idiom acquisition. Idiomatic meanings may not be learned adequately unless special attention is given to their semantic properties. Thus, acquiring L2 idioms would benefit from explicit focus during learning and teaching. L2 learners should be made aware that some MWS carry idiomatic meanings and be encouraged to interpret and use these meanings without relying heavily on their alternative literal interpretations.

Conclusion

This study investigated the processing of different types of MWSs in both L1 and L2 speakers. The findings confirmed that both L1 and L2 speakers processed collocations more quickly than novel word combinations, supporting the view that phrase-level frequency and familiarity play a crucial role in MWS processing. Additionally, L1 speakers exhibited a processing advantage for figuratively used idioms over collocations, even when distributional factors were controlled. Overall, the present findings suggest that MWS processing is best characterized not as a categorical distinction between idiomatic and compositional processing, nor as a qualitative difference between L1 and L2 systems, but as a graded architecture in which the relative contribution of direct retrieval and compositional computation varies continuously. This gradient is shaped by (a) the degree of semantic conventionality of the expression and (b) the amount of language experience available to the speaker.

Replication package. Replication data and materials for this article can be found at https://osf.io/fx3mr/.

Competing interests. The authors declare no competing interests.

Downloaded from https://www.cambridge.org/core. 26 Oct 2026 at 14:05:28, subject to the Cambridge Core terms of use.

References

Arnon, I., McCauley, S. M., & Christiansen, M. H. (2017). Digging up the building blocks of language: Age-of-acquisition effects for multiword phrases. *Journal of Memory and Language*, **92**, 265–280. https://doi.org/10.1016/j.jml.2016.07.004

Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language*, **62**(1), 67–82. https://doi.org/10.1016/j.jml.2009.09.005

Ayto, J. (2009). *Oxford dictionary of English idioms*. Oxford University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, **59**(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, **68**(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, **67**, 1–48. https://doi.org/10.18637/jss.v067.i01

Beck, S. D., & Weber, A. (2016). Bilingual and monolingual idiom processing is cut from the same cloth: The role of the L1 in literal and figurative meaning activation. *Frontiers in Psychology*, **7**, 1350. https://doi.org/10.3389/fpsyg.2016.01350

Bobrow, S. A., & Bell, S. M. (1973). On catching on to idiomatic expressions. *Memory & Cognition*, **1**(3), 343–346. https://doi.org/10.3758/BF03198118

Boland, J. E., Tanenhaus, M. K., & Garnsey, S. M. (1990). Evidence for the immediate use of verb control information in sentence processing. *Journal of Memory and Language*, **29**(4), 413–432. https://doi.org/10.1016/0749-596X(90)90064-7

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition*, **1**(1), 9-9. https://doi.org/10.5334/joc.10

Bybee, J. (1998). The Emergent Lexicon. In M. C. Gruber, D. Higgins, K. S. Olson, & T. Wysocki (Eds.), *CLS 34: The panels* (Vol. 2, pp. 421–435). Chicago Linguistic Society.

Cacciari, C., & Corradini, P. (2015). Literal analysis and idiom retrieval in ambiguous idioms processing: A reading-time study. *Journal of Cognitive Psychology*, **27**(7), 797–811. https://doi.org/10.1080/20445911.2015.1049178

Cacciari, C., & Tabossi, P. (1988). The comprehension of idioms. *Journal of Memory and Language*, **27**(6), 668–683. https://doi.org/10.1016/0749-596X(88)90014-9

Caillies, S., & Butcher, K. (2007). Processing of idiomatic expressions: Evidence for a new hybrid view. *Metaphor and Symbol*, **22**(1), 79–108. https://doi.org/10.1080/10926480709336754

Carrol, G., & Conklin, K. (2014). Getting your wires crossed: Evidence for fast processing of L1 idioms in an L2. *Bilingualism: Language and Cognition*, **17**(4), 784–797. https://doi.org/10.1017/S1366728913000795

Carrol, G., & Conklin, K. (2017). Cross language lexical priming extends to formulaic units: Evidence from eye-tracking suggests that this idea 'has legs'. *Bilingualism: Language and Cognition*, **20**(2), 299–317. https://doi.org/10.1017/S1366728915000103

Carrol, G., & Conklin, K. (2020). Is all formulaic language created equal? Unpacking the processing advantage for different types of formulaic sequences. *Language and Speech*, **63**(1), 95–122. https://doi.org/10.1177/0023830918823230

Carrol, G., Conklin, K., & Gyllstad, H. (2016). Found in translation: The influence of the L1 on the reading of idioms in a L2. *Studies in Second Language Acquisition*, **38**(3), 403–443. https://doi.org/10.1017/S0272263115000492

Carrol, G., & Segaert, K. (2024). As easy as cake or a piece of pie? Processing idiom variation and the contribution of individual cognitive differences. *Memory & Cognition*, **52**(2), 334–351. https://doi.org/10.3758/s13421-023-01463-x

Chen, C., & Truscott, J. (2010). The effects of repetition and L1 lexicalization on incidental vocabulary acquisition. *Applied Linguistics*, **31**(5), 693–713. https://doi.org/10.1093/applin/amq031

Christiansen, M. H., & Arnon, I. (2017). More than words: The role of multiword sequences in language learning and use. *Topics in Cognitive Science*, **9**(3), 542–551. https://doi.org/10.1111/tops.12274

Cieślicka, A. (2006). Literal salience in on-line processing of idiomatic expressions by second language learners. *Second Language Research*, **22**(2), 115–144. https://doi.org/10.1191/0267658306sr263oa

Downloaded from https://www.cambridge.org/core. 26 Oct 2026 at 14:05:28, subject to the Cambridge Core terms of use.

Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than non-formulaic language by native and nonnative speakers? *Applied Linguistics*, **29**(1), 72–89. https://doi.org/10.1093/applin/amm022

Cutting, J. C., & Bock, K. (1997). That's the way the cookie bounces: Syntactic and semantic components of experimentally elicited idiom blends. *Memory & Cognition*, **25**(1), 57–71. https://doi.org/10.3758/BF03197285

DeBruine, L. M., & Barr, D. J. (2021). Understanding mixed-effects models through data simulation. *Advances in Methods and Practices in Psychological Science*, **4**(1), 2515245920965119. https://doi.org/10.1177/2515245920965119

Diependaele, K., Lemhöfer, K., & Brysbaert, M. (2013). The word frequency effect in first-and second-language word recognition: A lexical entrenchment account. *Quarterly Journal of Experimental Psychology*, **66**(5), 843–863. https://doi.org/10.1080/17470218.2012.720994

Durrant, P., & Doherty, A. (2010). Are high-frequency collocations psychologically real? Investigating the thesis of collocational priming. *Corpus Linguistics and Linguistic Theory*, **6**(2), 125–155. https://doi.org/10.1515/cllt.2010.006

Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text & Talk*, **20**(1), 29–62. https://doi.org/10.1515/text.1.2000.20.1.29

Ferreira, F., & Patson, N. D. (2007). The "good enough" approach to language comprehension. *Language and Linguistics Compass*, **1**(1–2), 71–83. https://doi.org/10.1111/j.1749-818X.2007.00007.x

Gibbs, R. W. (1980). Spilling the beans on understanding and memory for idioms in conversation. *Memory & Cognition*, **8**(2), 149–156. https://doi.org/10.3758/BF03213418

Gibbs, R. W., & Nayak, N. P. (1989). Psycholinguistic studies on the syntactic behavior of idioms. *Cognitive Psychology*, **21**(1), 100–138. https://doi.org/10.1016/0010-0285(89)90004-2

Gibbs, R. W., Nayak, N. P., & Cutting, C. (1989). How to kick the bucket and not decompose: Analyzability and idiom processing. *Journal of Memory and Language*, **28**(5), 576–593. https://doi.org/10.1016/0749-596X(89)90014-4

Gibson, E., Pearlmutter, N., Canseco-Gonzalez, E., & Hickok, G. (1996). Recency preference in the human sentence processing mechanism. *Cognition*, **59**(1), 23–59. https://doi.org/10.1016/0010-0277(95)00687-7

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

Hamblin, J. L., & Gibbs, R. W. (1999). Why you can't kick the bucket as you slowly die: Verbs in idiom comprehension. *Journal of Psycholinguistic Research*, **28**(1), 25–39. https://doi.org/10.1023/A:1023235403250

Howarth, P. (1998). Phraseology and second language proficiency. *Applied Linguistics*, **19**(1), 24–44. https://doi.org/10.1093/applin/19.1.24

Hubers, F., Cucchiarini, C., & Strik, H. (2020). Second language learner intuitions of idiom properties: What do they tell us about L2 idiom knowledge and acquisition? *Lingua*, **246**, 102940. https://doi.org/10.1016/j.lingua.2020.102940

Hunston, S. (2022). *Corpora in applied linguistics* (2nd ed.). Cambridge University Press.

Jolsvai, H., McCauley, S. M., & Christiansen, M. H. (2020). Meaningfulness beats frequency in multiword chunk processing. *Cognitive Science*, **44**(10), e12885. https://doi.org/10.1111/cogs.12885

Kallens, P., & Christiansen, M. H. (2022). Models of language and multiword expressions. *Frontiers in Artificial Intelligence*, **5**, 781962. https://doi.org/10.3389/frai.2022.781962

Kuiper, K., van Egmond, M.-E., Kempen, G., & Sprenger, S. (2007). Slipping on superlemmas: Multi-word lexical items in speech production. *The Mental Lexicon*, **2**(3), 313–357. https://doi.org/10.1075/ml.2.3.03kui

Kuperman, V., & Van Dyke, J. A. (2013). Reassessing word frequency as a determinant of word recognition for skilled and unskilled readers. *Journal of Experimental Psychology: Human Perception and Performance*, **39**, 802–823. https://doi.org/10.1037/a0030859

Kyriacou, M., Conklin, K., & Thompson, D. (2020). Passivizability of idioms: Has the wrong tree been barked up? *Language and Speech*, **63**(2), 404–435. https://doi.org/10.1177/0023830919847691

Lenth, R. V., Bolker, B., Buerkner, P., Giné-Vázquez, I., Herve, M., Jung, M., Love, J., Miguez, F., Piaskowski, J., Riebl, H., & Singmann, H. (2024). emmeans: Estimated Marginal Means, aka Least-Squares Means (Version 1.10.3) [Computer software]. https://cran.r-project.org/web/packages/emmeans/index.html

Downloaded from https://www.cambridge.org/core. 26 Oct 2026 at 14:05:28, subject to the Cambridge Core terms of use.

Li, H., Paterson, K. B., Warrington, K. L., & Wang, X. (2022). Insights into the processing of collocations during L2 English reading: Evidence from eye movements. *Frontiers in Psychology*, **13**, 845590. https://doi.org/10.3389/fpsyg.2022.845590

Li, H., Warrington, K. L., Pagán, A., Paterson, K. B., & Wang, X. (2021). Independent effects of collocation strength and contextual predictability on eye movements in reading. *Language, Cognition and Neuroscience*, **36**(8), 1001–1009. https://doi.org/10.1080/23273798.2021.1922726

Li, W., Bassetti, B., & Frisson, S. (2025). Changes in word order do not eliminate the collocation advantage: An eye-tracking study of L1 and L2 speakers. *Bilingualism: Language and Cognition*, **29**(1), 92–106. https://doi.org/10.1017/S1366728925000057

Libben, M. R., & Titone, D. A. (2008). The multidetermined nature of idiom processing. *Memory & Cognition*, **36**(6), 1103–1121. https://doi.org/10.3758/MC.36.6.1103

Macis, M., & Schmitt, N. (2017). The figurative and polysemous nature of collocations and their place in ELT. *ELT Journal*, **71**(1), 50–59. https://doi.org/10.1093/elt/ccw044

Milburn, E., Vulchanova, M., & Vulchanov, V. (2021). Collocational frequency and context effects on idiom processing in advanced L2 speakers. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, **75**, 169–174. https://doi.org/10.1037/cep0000214

Milburn, E., & Warren, T. (2019). Idioms show effects of meaning relatedness and dominance similar to those seen for ambiguous words. *Psychonomic Bulletin & Review*, **26**(2), 591–598. https://doi.org/10.3758/s13423-019-01589-7

Moon, R. (1998). *Fixed expressions and idioms in English: A corpus-based approach*. Oxford University Press.

Öksüz, D., Brezina, V., Monaghan, P., & Rebuschat, P. (2024). Individual word and phrase frequency effects in collocational processing: Evidence from typologically different languages, English and Turkish. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **50**(8), 1287–1314. https://doi.org/10.1037/xlm0001324

Öksüz, D., Brezina, V., & Rebuschat, P. (2021). Collocational processing in L1 and L2: The effects of word frequency, collocational frequency, and association. *Language Learning*, **71**(1), 55–98. https://doi.org/10.1111/lang.12427

Ortony, A., Schallert, D. L., Reynolds, R. E., & Antos, S. J. (1978). Interpreting metaphors and idioms: Some effects of context on comprehension. *Journal of Verbal Learning and Verbal Behavior*, **17**(4), 465–477. https://doi.org/10.1016/S0022-5371(78)90283-9

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. 191–226). Longman.

Pinker, S. (1999). *Words and rules: The ingredients of language*. Basic Books.

R Core Team. (2022). R: A language and environment for statistical (Version 4.2.1) [Computer software]. https://www.R-project.org/

Rebuschat, P., & Williams, J. N. (2012). Implicit and explicit knowledge in second language acquisition. *Applied Psycholinguistics*, **33**(4), 829–856. https://doi.org/10.1017/S0142716411000580

Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *WIREs Cognitive Science*, **1**(6), 906–914. https://doi.org/10.1002/wcs.78

Saffran, J. R., & Kirkham, N. Z. (2018). Infant statistical learning. *Annual Review of Psychology*, **69**(1), 181–203. https://doi.org/10.1146/annurev-psych-122216-011805

Sag, I. A., Baldwin, T., Bond, F., Copestake, A., & Flickinger, D. (2002). Multiword expressions: A pain in the neck for NLP. *Computational Linguistics and Intelligent Text Processing*, 1–15.

Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing*, **18**(1), 55–88.

Senaldi, M. S. G., & Titone, D. (2024). Idiom meaning selection following a prior context: Eye movement evidence of L1 direct retrieval and L2 compositional assembly. Discourse Processes. https://www.tandfonline.com/doi/full/10.1080/0163853X.2024.2311637

Senaldi, M. S. G., & Titone, D. A. (2022). Less direct, more analytical: eye-movement measures of L2 idiom reading. *Languages*, **7**(2), Article 2. https://doi.org/10.3390/languages7020091

Shi, J., Peng, G., & Li, D. (2023). Figurativeness matters in the second language processing of collocations: Evidence from a self-paced reading experiment. *Language Learning*, **73**(1), 47–83. https://doi.org/10.1111/lang.12516

Downloaded from https://www.cambridge.org/core. 26 Oct 2026 at 14:05:28, subject to the Cambridge Core terms of use.

Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford University Press.

Singh, R., Fedorenko, E., Mahowald, K., & Gibson, E. (2016). Accommodating presuppositions is inappropriate in implausible contexts. *Cognitive Science*, **40**(3), 607–634. https://doi.org/10.1111/cogs.12260

Siyanova, A., & Schmitt, N. (2008). L2 learner production and processing of collocation: A multi-study perspective. *The Canadian Modern Language Review*, **64**(3), 429–458. https://doi.org/10.3138/cmlr.64.3.429

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**(2), 251–272. https://doi.org/10.1177/0267658310382068

Siyanova-Chanturia, A., & Martinez, R. (2015). The idiom principle revisited. *Applied Linguistics*, **36**(5), 549–569. https://doi.org/10.1093/applin/amt054

Sonbul, S. (2015). Fatal mistake, awful mistake, or extreme mistake? Frequency effects on off-line/on-line collocational processing. *Bilingualism: Language and Cognition*, **18**(3), 419–437. https://doi.org/10.1017/S1366728914000674

Sprenger, S. A., Levelt, W. J. M., & Kempen, G. (2006). Lexical access during the production of idiomatic phrases. *Journal of Memory and Language*, **54**(2), 161–184. https://doi.org/10.1016/j.jml.2005.11.001

Swinney, D. A., & Cutler, A. (1979). The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behavior*, **18**(5), 523–534. https://doi.org/10.1016/S0022-5371(79)90284-6

Tabossi, P., Fanari, R., & Wolf, K. (2009). Why are idioms recognized fast? *Memory & Cognition*, **37**(4), 529–540. https://doi.org/10.3758/MC.37.4.529

Titone, D., Lovseth, K., Kasparian, K., & Tiv, M. (2019). Are figurative interpretations of idioms directly retrieved, compositionally built, or both? Evidence from eye movement measures of reading. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, **73**(4), 216–230. https://doi.org/10.1037/cep0000175

Titone, D. A., & Connine, C. M. (1994). Descriptive norms for 171 idiomatic expressions: Familiarity, compositionality, predictability, and literality. *Metaphor and Symbolic Activity*, **9**(4), 247–270. https://doi.org/10.1207/s15327868ms0904_1

Titone, D. A., & Libben, M. (2014). Time-dependent effects of decomposability, familiarity and literal plausibility on idiom priming: A cross-modal priming investigation. *The Mental Lexicon*, **9**(3), 473–496. https://doi.org/10.1075/ml.9.3.05tit

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Harvard university press.

Underwood, G., Schmitt, N., & Galpin, A. (2004). The Eyes Have It: An Eye-Movement Study into the Processing of Formulaic Sequences. In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing, and use* (pp. 53–172). John Benjamins.

Van Lancker Sidtis, D. (2012). Formulaic language and language disorders. *Annual Review of Applied Linguistics*, **32**, 62–80. https://doi.org/10.1017/S0267190512000104

Vilkaitė, L. (2016). Are nonadjacent collocations processed faster? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **42**, 1632–1642. https://doi.org/10.1037/xlm0000259

Vilkaitė, L., & Schmitt, N. (2019). Reading collocations in an L2: Do collocation processing benefits extend to non-adjacent collocations? *Applied Linguistics*, **40**(2), 329–354. https://doi.org/10.1093/applin/amx030

Vilkaitė-Lozdienė, L., & Conklin, K. (2021). Word order effect in collocation processing. *The Mental Lexicon*, **16**(2–3), 362–396. https://doi.org/10.1075/ml.20022.vil

Wang, T., & Saffran, J. R. (2014). Statistical learning of a tonal language: The influence of bilingualism and previous linguistic experience. *Frontiers in Psychology*, **5**, 953. https://doi.org/10.3389/fpsyg.2014.00953

Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, **28**(1), 46–65. https://doi.org/10.1093/applin/aml048

Webb, S., Newton, J., & Chang, A. (2013). Incidental learning of collocation. *Language Learning*, **63**(1), 91–120. https://doi.org/10.1111/j.1467-9922.2012.00729.x

Wolter, B., & Gyllstad, H. (2011). Collocational links in the L2 mental lexicon and the influence of L1 intralexical knowledge. *Applied Linguistics*, **32**(4), 430–449. https://doi.org/10.1093/applin/amr011

Wolter, B., & Gyllstad, H. (2013). Frequency of input and L2 collocational processing a comparison of congruent and incongruent collocation. *Studies in Second Language Acquisition*, **35**(3), 451–482. https://doi.org/10.1017/S0272263113000107

Wolter, B., & Yamashita, J. (2015). Processing collocations in a second language: A case of first language activation? *Applied Psycholinguistics*, **36**(5), 1193–1221. https://doi.org/10.1017/S0142716414000113

Downloaded from https://www.cambridge.org/core. 26 Oct 2026 at 14:05:28, subject to the Cambridge Core terms of use.

Wolter, B., & Yamashita, J. (2018). Word frequency, collocational frequency, L1 congruency, and proficiency in L2 collocational processing: What accounts for L2 performance? *Studies in Second Language Acquisition*, **40**(2), 395–416. https://doi.org/10.1017/S0272263117000237

Wray, A. (2002). Formulaic language in computer-supported communication: Theory meets reality. *Language Awareness*, **11**(2), 114–131. https://doi.org/10.1080/09658410208667050

Yi, W., & Zhong, Y. (2024). The processing advantage of multiword sequences: A meta-analysis. *Studies in Second Language Acquisition*, **46**(2), 427–452. https://doi.org/10.1017/S0272263123000542

Cite this article: Li, W., Bassetti, B., & Frisson, S. (2026). The uneven advantage of multiword sequences: How idioms and collocations are processed in L1 and L2. *Applied Psycholinguistics*. https://doi.org/10.1017/S0142716426100939

Downloaded from https://www.cambridge.org/core. 26 Oct 2026 at 14:05:28, subject to the Cambridge Core terms of use.