



Research Article  

Cite this article: Li, W., Bassetti, B. and 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* 1–15. <https://doi.org/10.1017/S1366728925000057>

Received: 03 June 2023
Revised: 12 December 2024
Accepted: 14 December 2024

Keywords:
collocations; variation collocations; multi-word sequences; bilingualism; eye-tracking

Corresponding author:
Wanyin Li;
Email: wanyin.li.words@gmail.com

  This research article was awarded Open Data and Open Materials badges for transparent practices. See the Data Availability Statement for details.

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

Changes in word order do not eliminate the collocation advantage: An eye-tracking study of L1 and L2 speakers

Wanyin Li¹ , Bene Bassetti^{1,2} and Steven Frisson³

¹School of Education, University of Birmingham, Birmingham, B15 2TT, United Kingdom; ²Department of Education and Humanities, University of Modena and Reggio Emilia, Viale Timavo 93, 42121 Reggio Emilia, Italy and ³School of Psychology, University of Birmingham, Birmingham, B15 2TT, United Kingdom

Abstract

Collocations, defined as sequences of frequently co-occurring words, show a processing advantage over novel word combinations in both L1 and L2 speakers. This collocation advantage is mainly observed for canonical configurations (e.g., *provide information*), but collocations can also occur in variation configurations (e.g., *provide some of the information*). Variation collocations still show a processing advantage in L1 speakers, but generally not in L2 speakers. The present eye-tracking-while-reading experiment investigated word order variation by passivising collocations (e.g., *information was provided*) in L1 and advanced L2 speakers of English. Altering word order did not eliminate the collocation advantage in either L1 or L2 speakers. The collocation effect was independent of contextual predictability and modulated by L2 proficiency. Results support the view that collocations are stored and retrieved via semantic representation rather than as holistic form chunks and that collocation processing does not qualitatively differ between L1 and advanced L2 speakers.

Highlights

- Changing the word order of L1 collocations did not disrupt the processing advantage
- L2 speakers also maintained the processing advantage in reversed collocations
- The collocation advantage was independent of contextual predictability

1. Introduction

Collocations have attracted much attention among researchers investigating language processing in both first (L1) and second (L2) languages due to consistent evidence of a ‘collocation advantage’. This refers to the phenomenon where collocations (e.g., *provide information*) are processed faster than novel¹ word combinations (e.g., *compare information*). The collocation advantage suggests that the frequency effect on language processing extends beyond single words and applies to multi-word sequences (MWSs) as well. While such evidence aligns with current views that MWSs – including collocations, idioms, and lexical bundles – are fundamental building blocks of language (Arnon et al., 2017; Arnon & Snider, 2010), this view of MWSs contrasts with earlier research, which for a long time focussed on single words (Pinker, 1991; Pinker & Prince, 1988).

Despite the growing body of research on collocations, there is limited understanding of how variations in collocation configurations – such as word order changes – affect processing, particularly in natural reading contexts. Prior research on collocations has largely concentrated on their canonical forms, revealing consistent processing advantages for L1 and proficient L2 speakers alike (e.g., Li et al., 2022; Öksüz et al., 2021; Sonbul, 2015). However, studies addressing variation in collocations remain sparse, with mixed findings on the effects of non-adjacency (Vilkaitė, 2016; Vilkaitė & Schmitt, 2019), morphological variations (Vilkaitė-Lozdienė, 2022), and word order changes (Vilkaitė-Lozdienė & Conklin, 2021). Furthermore, while the flexibility of collocations makes them a critical area of investigation distinct from other fixed expressions, studies on their non-canonical forms have primarily examined L1 speakers, leaving gaps in understanding L2 processing. This study addresses these limitations by exploring the effects of word order changes in L1 and L2 collocations using eye-tracking during reading, a method that provides fine-grained insights into both early and later stages of processing.

¹To align with the literature (e.g., Kyriacou et al., 2020; Vilkaitė-Lozdienė, 2022; Vilkaitė-Lozdienė & Conklin, 2021), we used the term ‘novel’ to refer to the baseline. However, this does not imply that a novel word combination is entirely new. Rather, ‘novel’ refers to word combinations less probable than highly frequent collocations.

1.1. Defining and identifying collocations

We define collocations as sequences of words that co-occur more frequently than we expect by chance (Biber et al., 1999); for example, *provide information* is more frequent than novel sequences of words such as *compare information*. Based on this definition, we identify collocations using the frequency-based approach (Sinclair, 1991), whereby collocations are extracted from a corpus using association measures that measure the strength of the co-occurrence between two words (Hunston, 2022). This approach is generally adopted in previous eye-tracking studies of collocation processing (e.g., Li et al., 2021; Sonbul, 2015; Vilkaitė, 2016; Vilkaitė-Lozdienė 2022). Among the association measures that can be used to identify collocations, following Vilkaitė-Lozdienė (2022), we adopted not only the frequency of the phrase but also the Mutual Information (MI) score², a common measure of exclusivity. A higher MI score indicates that the words in a collocation are more exclusively associated and less likely to form other collocations. We included only phrases with an MI score larger than 3 (Hunston, 2022). To avoid including idioms, we excluded phrases that were listed in the Oxford Dictionary of English Idioms (Ayto, 2009).

The frequency-based approach we adopted is not the only one, and different approaches identify partly different sets of collocations. The alternative approach, the phraseology-based method, emphasises the role of semantic specification and restrictedness in lexical compositionality (Cowie, 1998; Howarth, 1998; Nesselhauf, 2003), with expressions becoming increasingly more restricted from free combinations to collocations to idioms. Thus, collocations identified with this method are generally neither fully compositional nor fully semantically transparent (e.g., *take pictures*, *draw conclusions*). Using a phraseology-based approach, Gyllstad and Wolter (2016) found that increased semantic restrictedness led to increased collocation processing effort for both L1 and L2 speakers (Gyllstad & Wolter, 2016).

1.2. Collocation processing in L1

Studies using an adapted lexical decision task found that L1 speakers process collocations (e.g., *provide information*) faster than novel word combinations (e.g., *compare information*) (e.g., Durrant & Doherty, 2010; Siyanova & Schmitt, 2008; Wolter & Gyllstad, 2011, 2013; Wolter & Yamashita, 2015, 2018). This collocation advantage cannot be fully explained by associative priming (Durrant & Doherty, 2010), even though collocation components can be associated, as shown by word association norms (e.g., the Edinburgh Association Thesaurus, Kiss et al., 1973).

Eye-tracking studies identified other factors affecting the collocation effect. Tasks that involve judgements, such as the ones above, do not reflect natural reading and only measure the endpoint of processing. Eye-tracking provides a more ecological approach to researching collocation processing. It also provides measures tapping into earlier stages of processing (e.g., first fixation duration, gaze duration) and those that capture reanalysis and deeper processing (e.g., total reading time, fixation counts) (Clifton et al., 2007; Holmqvist et al., 2011; Rayner, 1998).

Eye-tracking studies generally use different approaches to identify collocations. Early studies (Frisson et al., 2005; McDonald & Shillcock, 2003b, 2003a) used transitional probabilities (TP, the statistical likelihood that a word will precede or follow another

word). More recent studies (Li et al., 2021; Sonbul, 2015; Vilkaitė, 2016; Vilkaitė-Lozdienė 2022) use collocational frequency and MI scores that take into account a more flexible collocation window. Both methods are frequency-based and use quantitative evidence of word co-occurrence in corpora but result in different selections of collocations (Gablasova et al., 2017). To illustrate, word pairs that were selected based on an MI score of 3 (Vilkaitė, 2016) had much higher phrase-level frequency and TP than those selected based on TP (Frisson et al., 2005) (360 versus 60, and .0127 versus .00677, respectively)³.

Early eye-tracking studies using TP found that collocations (high-TP) are read faster than control (low-TP) sequences in L1 speakers, but also that this advantage may result from contextual predictability. McDonald and Shillcock (2003a, 2003b) found that high-TP verb-noun pairs (e.g., *accept defeat*) had shorter reading times than low-TP ones (e.g., *accept losses*) in L1 speakers. This appeared in early eye movement measures, suggesting that TP affects immediate processing. However, Frisson et al. (2005) extended this research by introducing a contextual manipulation that made the high-TP and low-TP verb-noun pairs either predictable or not from the preceding context and found strong early effects of contextual predictability, but no independent effect of TP. Frisson et al. suggested that contextual predictability rather than TP affects reading and that TP is one of the sources readers can use to guide prediction. Hence, it can be challenging to disentangle contextual predictability and TP because contextual predictability, as measured by cloze probabilities (see below), likely captures some degree of TP. This raises questions about whether collocation effects are driven by contextual predictability.

Recent eye-tracking research on collocations has adopted both phrase-level frequency and MI as the selection method and consistently found that collocations identified in this way are read faster than novel word combinations in L1 speakers (Li et al., 2021; Sonbul, 2015; Vilkaitė, 2016; Vilkaitė-Lozdienė 2022). Furthermore, Li et al. (2021) used a 2x2 design (collocation strength: strong collocations versus weak collocations; context: predictable versus neutral context) with L1 speakers. Collocation strength and predictability effects were observed without interaction, indicating that the collocation strength effect found in both early and late eye-tracking measures did not depend on contextual predictability.

These studies found that collocations are read faster than non-collocations and this effect appears in early eye-tracking measures (e.g., first-run reading times). However, it seems premature to accept this conclusion due to difficulty in controlling confounds. For example, Li et al. (2021), discussed above, operationalised contextual predictability as cloze probability (Taylor, 1953). The widely used cloze probability test estimates predictability by asking participants to provide the next word(s) to continue a sentence fragment. Li et al. asked participants to fill in a word for an incomplete sentence up to the start of the collocation (e.g., *My friend and I made a _____*), providing a predictability value for the first word (*fatal*) of the collocation (*fatal mistake*). However, no cloze test was done for the second word of the MWS (which can be estimated using the same task but including the first word, e.g., *My friend and I made a fatal _____*). If, as argued by Frisson et al. (2005), TP can affect cloze, the predictability of the second word is likely higher for collocations than in matched free sequences, which could have driven the early collocation effect found in Li et al. (2021).

²MI scores operate on a scale without a minimum and maximum (Gablasova et al., 2017).

³Phrase-level frequencies were extracted from the BNC and are expressed as raw occurrences per 100 million. TP was calculated using the raw number of occurrences.

Most studies mentioned above focused on collocations in their canonical forms. However, it is also important to investigate whether the collocation advantage extends to non-canonical forms, as modified collocations are common. This could shed light on whether collocations are stored and processed as fixed units with distinct forms and meanings or accessed and produced based on their meaning, even when the form varies.

1.3. Processing non-canonical collocations in L1

Multi-word sequences, including collocations, have a canonical form (Moon, 1998) but can undergo some degree of form variation while retaining their meaning. The form of collocations can be modified in different ways: (1) inserting words between the component words, making them non-adjacent (e.g., *provide information* versus *provide some of the information*), (2) varying morphological forms (*provide information* versus *providing/provided information*), and (3) varying word order (*provide information* versus *information was provided*).

Few studies have looked at the effect of variation on MWSs processing in L1 speakers, and the evidence on whether modified MWSs retain an online processing advantage is inconsistent. Molinaro et al. (2013) combined neural and behavioural measures and modified lexical bundles (e.g., *in the hands of*) by inserting a pre-modifying adjective (*capable*) before the first noun (*hands*) (e.g., *in the capable hands of*). The self-paced reading results showed longer reading times for both the noun (*hands*) and the following word (*of*) in the variation condition, compared to the canonical forms. The longer RTs in the variation condition suggest that extra processing is required when the canonical form is disrupted, or additional effort is needed to integrate the inserted word into the context. In contrast, there was a reduced N400⁴ amplitude for the noun (*hands*) in the variation condition compared to the lexical bundles, indicating easier semantic processing and contextual integration. The reduced N400 may show that inserting an adjective does not disrupt lexical bundle processing. However, a Left Anterior Negativity⁵ (LAN) was found for the noun that followed a variation lexical bundle (e.g., *restorers in in the (capable) hands of restorers*), suggesting additional difficulty in processing and integrating non-canonical MWSs. In short, the mixed findings of the study by Molinaro et al. do not provide conclusive evidence for a processing (dis)advantage for variation MWSs.

Conversely, two eye-tracking studies with L1 speakers showed that morphological variation and non-adjacent collocations can retain a processing advantage. Participants consistently processed collocations faster than novel word combinations in both early and late eye-tracking measures with both adjacent and non-adjacent verb-noun collocations (e.g., *provide information* and *provide some of the information*; Vilkaitė, 2016). Furthermore, Vilkaitė-Lozdienė (2022) found that collocations retained their processing advantage when their morphological forms were changed. Native Lithuanian readers were presented with stories containing verb-object collocations and novel word combinations, each in three morphological forms: the canonical form, namely infinitive + accusative (e.g., *švęsti pergale*, 'to celebrate the victory'), past tense third person + accusative (e.g., *šventė pergale*, 'celebrated the victory'), and attributive passive participle (e.g., *švęsta pergale*, 'the victory one celebrated').

⁴The N400 is closely related to the semantic processing of a word and the ease of integrating it into an unfolding sentence (Holcomb, 1988).

⁵LAN is often evoked by anomalies or increased working load (Kutas & Federmeier, 2000).

At the phrase level (i.e., whole-phrase reading times), collocations were read faster than controls across morphological form variations. The collocation advantage was found only in late (i.e., total reading times), but not early, eye-tracking measures. Additionally, collocations in their canonical form were read the fastest, while past tense collocations were read the slowest, though still faster than their corresponding controls. At the final word level (i.e., reading times for the final word of collocations), both early and late measures revealed faster reading times for all collocation forms, and collocation variations were fixated less often than controls.

To our knowledge, there have been no eye-tracking-during-reading studies of word order changes on collocation processing. What evidence we have comes from a judgment task study of collocations (Vilkaitė-Lozdienė & Conklin, 2021) and an eye-tracking study of idioms (Kyriacou et al., 2020). In a primed lexical decision task, Vilkaitė-Lozdienė and Conklin (2021) compared the effects of forward and backward priming on collocation processing in natives of English and Lithuanian, the latter a language with higher flexibility of word order. In the forward condition, the prime (e.g., *attract*) was the first word of a collocation (e.g., *attract attention*), and the target (e.g., *attention*) was the second word of the collocation. In the backward condition, the prime (e.g., *attention*) was the second word of a collocation, and the target (e.g., *attract*) was the first word of the collocation. Lithuanians showed facilitative priming compared to control words in both conditions. English natives showed a facilitative priming effect in the forward but not in the backward condition due to the low flexibility of word order in English. In an eye-tracking study examined word order change for idioms in L1 English speakers, using passive constructions (Kyriacou et al., 2020). They compared canonical idioms (e.g., *he kicked the bucket*) and idioms in their reversed forms (*the bucket was kicked*) to their counterpart controls (e.g., *he kicked the apple* versus *the apple was kicked*). Late eye-tracking measures for both the phrase-level and word-level regions showed that word order (canonical or reversed) and idiomaticity (idiom or control) were significant predictors, and there was no interaction: idioms were read faster than controls both in canonical and reversed word order, and canonical idioms were read faster than reversed ones. These findings suggest that reversed idioms retain the figurative meaning associated with their canonical form and order. The present study differs from Kyriacou et al. (2020) by focusing on collocations rather than idioms. Collocations are more flexible and less rigid in structure than idioms. Altering their word order can disrupt meaning to varying extents, potentially leading to processing patterns that differ from the more conventional idioms. Furthermore, the current study explores how L2 speakers process non-canonical collocations, investigating whether their approach aligns with that of L1 speakers.

In conclusion, research on modified MWSs in L1 speakers is limited and shows mixed results. Some studies indicate increased processing effort with disrupted canonical forms, while others show a retained advantage with non-adjacency, morphological form variations, and word order changes. The situation is even more challenging for L2 speakers, with only one study on processing modified MWSs.

1.4. Processing collocations in L2

Although the studies above discussed the collocation processing advantage in L1 speakers, a collocation advantage is also consistently found in proficient L2 speakers of English (Li et al., 2022; Öksüz et al., 2021; Siyanova & Schmitt, 2008; Sonbul, 2015; Vilkaitė

& Schmitt, 2019; Wolter & Gyllstad, 2011, 2013). It appears that phrase-level frequency affects both L1 and L2 speakers. For example, L2 speakers judge more frequent collocations faster than less frequent ones, similar to L1 speakers (Siyanova & Schmitt, 2008), and show a processing advantage for more frequent L2 collocations in early eye-tracking measures (Sonbul, 2015). Crucially, all these studies found similarities between L1 and L2 speakers in the processing of canonical collocations.

To our knowledge, only one study has investigated the effects of variation on L2 collocation processing, finding no collocation advantage. Vilkaitė and Schmitt (2019) presented advanced L2 speakers with the canonical and modified (non-adjacent) collocations used by Vilkaitė (2016) with L1 speakers. The eye-tracking data showed no processing advantage for non-adjacent collocations in L2 speakers, though both L1 and L2 speakers showed an advantage for canonical collocations. However, these L2 speakers had various first languages, which could be a limitation because Wolter and Gyllstad (2011, 2013) found that the L1 influences L2 collocation processing. Specifically, L2 speakers responded faster to L2 collocations that were congruent (i.e., having a semantically equivalent collocation in the L1) compared with incongruent ones. To address the issue, the present study restricted the sample to a single L1 (Chinese) and focussed solely on congruent L2 collocations.

1.5. The present study

Both L1 and L2 speakers show a consistent advantage for canonical collocations, but research on non-canonical forms has inconsistent results. Furthermore, because research predominantly investigated L1 speakers, it remains unclear how L2 speakers process modified form collocations. The present study investigated how changes in word order affect collocation processing during sentence reading for both L1 and L2 speakers. Specifically, we examined whether collocations maintained their processing advantage when the word order was reversed (e.g., *she brought peace* -> *peace was brought*) and whether similar effects were observed in advanced L2 speakers. We investigated word order because it had been studied in idioms (Kyriacou, 2020) but not in collocations. We adopted eye-tracking, a more sensitive method than judgment tasks (as used in, e.g. Vilkaitė-Lozdiene & Conklin, 2021), and followed Kyriacou (2020) in changing word order using passive constructions. We controlled for context predictability in order to determine whether the collocation effect is distinct from contextual predictability and to identify when the collocation advantage emerges during reading – whether immediately or only at a later stage of processing – in both L1 and L2 readers. We tested advanced L2 speakers because only one eye-tracking study had investigated their modified collocation processing (Vilkaitė-Lozdiene, 2022). We tested L2 speakers with the same L1 (Chinese) to control for L1 congruency and measured L2 speakers' proficiency to investigate whether proficiency modulates the collocation effect.

2. Method

2.1. Participants

Thirty-nine Chinese speakers of L2 English (age: $M = 23.8$, $SD = 3.06$) and 40 native speakers of British English (age: $M = 22$, $SD = 3.93$) were paid £7 to participate (one additional L2 speaker was removed due to track loss). Because this was, to our knowledge, the first eye-tracking study investigating word order effects on collocation processing, it was difficult to conduct an appropriate

power analysis for mixed effects models. To achieve a reliable effect size, the present study had more observations (1920) than recent eye-tracking collocations studies (1536 in Li et al., 2021; 1120 in Vilkaitė, 2016) per group. L2 participants' English proficiency was estimated using the LexTALE (Lexical Test for Advanced Learners of English, Lemhöfer & Broersma, 2012) and Vocabulary Levels Test of 3 k, 5 k and 10 k (VLT, Schmitt et al., 2001). The vocabulary score was used as an indicator of general English proficiency. The average LexTALE score was 66.21 (out of 100) and the average VLT score was 53.82 (out of 90), indicating relatively high L2 proficiency. Both measurements, VLT and LexTALE, were included because some previous research on modification collocation used LexTALE (e.g., Vilkaitė-Lozdiene & Conklin, 2021) while others used VLT (e.g., Vilkaitė & Schmitt, 2019). Including both measures allowed for a comparison of the current results and characteristics of L2 participants with those from previous research. On average, L2 speakers started learning English at age 7.42 and spent 2.5 years in an English-speaking country. Participants' language background and usage data are summarised in [Supplementary materials Table S1](#). All participants had normal or corrected vision and no self-reported history of reading difficulties.

The participants gave informed consent. The experiment was conducted in accordance with the British Psychological Society ethical guidelines and was approved by the University of Birmingham's ethical committee (ERN_19-0936).

2.2. Materials

Materials consisted of 192 verb-noun phrases. There were 48 sets of four phrases, corresponding to four conditions. Items were embedded within sentences consisting of four sections (see [Figure 1](#)). The pre-target contexts and spill-over regions were the same across the four conditions. The end-of-sentence region was kept the same whenever possible, with some minor changes for naturalness (note that this region was not analysed). Pre-target contexts were as neutral and uninformative as possible, and collocations were placed in the middle of sentences to avoid line breaks and sentence wrap-up effects (Rayner et al., 2000).

Materials preparation. Initially, we selected frequent phrases with an MI score over 3. Target phrases contained words from the 2000 most frequent verb and noun lemmas in the British National Corpus (BNC). We selected only collocations and novel word combinations that were usable in the passive voice. To control for L1-L2 congruency effects, all collocations had a semantic equivalent in Chinese (i.e., a direct word-for-word translation, e.g., *bring peace*, 带来和平), as assessed by five Chinese native speakers. Seventy-nine collocation-novel word combination pairs were selected for testing.

To match the lengths of the areas of interest in the canonical and reversed conditions, a subject was added before canonical phrases (Kyriacou et al., 2020). For example, *he kicked the bucket* and *the bucket was kicked* both consist of four words. The tense was also kept the same (e.g., *they brought peace* -> *peace was brought*; *he had prevented damage* -> *damage had been prevented*).

Predictability. To assess predictability for final words in the context sentences, 79 sentence frames were created, into which a collocation and its corresponding novel word combination could be inserted interchangeably (see [Figure 1](#)). The sentence frames were split into four lists by a Latin square, with an item only appearing once in each list, and each list containing a different condition of the same item. Sixty British English participants (15 per list) provided written continuations for sentences truncated immediately before

Word order	Pre-target context	Target region	Spill-over region	End of sentence
Canonical	Margaret is not sure that	<u>they brought peace/meat</u>	to the	village.
Reversed		<u>peace/meat was brought</u>		

Figure 1. Example of a sentence with a collocation and its control in canonical and reversed order.
 Note: Collocations are underlined. Collocations and controls are separated by a slash.

the final word of the collocation or novel word combinations (e.g., *Margaret is not sure that they brought _____, Margaret is not sure that meat was _____, Margaret is not sure that peace was _____*).

After removing 31 pairs with high context predictability scores, 48 pairs were selected. Their predictability did not statistically differ between canonical collocations and canonical novel word combinations ($t(47) = 0.44, p = .66$), and between reversed collocations and reversed novel word combinations ($t(47) = 1.69, p = .09$). The mean cloze values for the final words of the 48 collocations (and their controls) were below or near 1%. Hence, the preceding contexts provided little information regarding the sentence continuation, and collocations and novel word combinations were equally non-predictable. Finally, another 15 native British speakers were asked to continue the selected 48 sentences truncated before the whole collocation/control (e.g., *Margaret is not sure that _____*). Predictability for these fragments was zero. Predictability scores are reported in [Supplementary materials Table S2](#).

Frequency and length. The 48 pairs of collocations and novel word combinations were closely matched for individual word length and word frequency (see [Supplementary materials Table S2](#)). All frequency information, including individual word frequency and phrasal-level frequency with different morphological forms, was obtained from the BNC using Sketch Engine (Kilgarriff et al., 2004). The verbs in each pair did not differ in lemma frequency ($t(47) = -0.08, p = .93$), past participle frequency ($t(47) = 0.19, p = .84$), and root frequency ($t(47) = -0.4, p = .64$). The nouns also did not differ in lemma frequency ($t(47) = -0.23, p = .81$) or form frequency ($t(47) = 1.50, p = .13$). The verbs', nouns', and items' lengths were also matched within each pair ($t(47) = 0.44, p = .66$; $t(47) = -0.57, p = .57$; $t(47) = -0.70, p = .49$).

Apparatus and procedure. Participants were tested individually in a quiet laboratory using an Eyelink 1000 SR Research eye-tracker (SR Research, Ontario, Canada) with a tower mount recording the right eye gaze location every millisecond during reading. All sentences were displayed on one or two lines, and the regions of interest never appeared at the end of a line. Sentences were presented in size 14 Courier New font as black text with white background. At the beginning of the experiment, a 5-point calibration procedure was carried out, with re-calibrations as needed. Participants were instructed to read silently for comprehension at a normal pace. At the start of each trial, a fixation cross appeared on the left of the screen. After participants fixated on the cross, the cross was replaced by a sentence. After reading a sentence, participants pressed the spacebar to remove it. To check that participants were reading for meaning, half of the sentences were followed by a comprehension question. Half of the questions required a “Yes” answer and half a “No” answer. The reading experiment took about 20 mins for L1 participants and 30 mins for L2 participants. Participants then completed a linguistic background questionnaire and the two vocabulary tests. This took around 15 mins.

Pre-processing of data. Prior to analyses, sentences with major track loss, for example, as a result of blinking, were removed (1% of

trials). Following software standard fixation cleaning procedures, short fixations below 80 ms were merged with nearby longer fixations. Fixations under 100 ms and above 1000 ms were trimmed. Participants generally had no difficulties answering comprehension questions, with an average accuracy of 91% (92% for L1 speakers, 90% for L2 speakers). Trials with incorrect comprehension answers were removed (5% of data). Fixation times more than 2.5 SDs from the mean for each participant were excluded as outliers (3% of data).

Measures. Following (Carroll and Conklin 2014a) recommendation and previous collocation studies (e.g., Li et al., 2021; Vilkaitė-Lozdienė, 2022; Vilkaitė, 2016), eye movement measures are reported for two regions: the whole MWS and its final word. We examined both because (1) the effect is more likely observed on the final word after encountering the rest of the MWS, and (2) the effect might occur both at the final word and when re-reading the MWS.

To capture both early and late processing, several reading measures were used. For the whole collocation region, early processing measures consisted of first run reading time (sum of fixations on a region before exiting the region for the first time); later processing measures included regression path reading time (sum of all fixation durations from entering an interest area until moving further in the text) and selective regression path reading time (sum of fixation durations from entering an interest area before moving further in the text, excluding fixations on previous regions); and comprehensive measures included total reading time (sum of all fixation durations in a region of interest) and total fixation count (sum of the number of fixations in a region of interest). For the final word analysis, we report first fixation reading time (first fixation on a word before exiting the region for the first time), first run reading time, regression path reading time, selective regression path reading time, total reading time, and total fixation count.

2.3. Statistical analysis

All analyses were carried out using R version 4.2.1 (R Core Team, 2022) with packages *lme4* (Bates et al., 2015) and *lmerTest* (Kuznetsova et al., 2020). Data were analysed with a (generalised) linear mixed effect model. Language-background was a between-group factor with two levels (L1 speakers versus L2 speakers). There were two within-group factors, each with two levels: phrase-type (novel word combinations versus collocations) and word-order (canonical versus reversed).

First, models were fitted with L1 and L2 speakers' data separately, in order to see how proficiency affected processing in L2 speakers and when collocation advantage showed up in L1 speakers. Then, models were constructed with both L1 and L2 speakers' data together to test for differences in collocation processing between L1 and L2 speakers. For each dependent variable and each region of interest, a core model included: phrase-type (sum coded: novel word combination: -1 versus collocation: 1), word-order (sum coded: reversed: -1 versus canonical: 1) and their interaction as fixed effects. 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 did not occur, “bobyqa” or “optimx” optimiser (Powell, 2009) was used to increase iterations, before simplifying random effects by removing interactions and then manipulation variables in the order of least variance explained until the model converged (Veldre & Andrews, 2014). Trial number was centred and entered into models as a covariate. For each dependent variable, we examined and analysed the entire MWSs region and the final word. Final words from reversed and canonical constructions were analysed separately because the final words were verbs in the passive constructions while nouns occupied the final position in the active constructions. In addition, separate analyses avoided comparing final words from different syntactic constructions. We also combined nouns from canonical and reversed constructions to run these analyses (See [Supplementary materials S1](#)).

For the L1 speakers, the full model included the core model and trial number. For the L2 speakers, vocabulary scores and the interactions with vocabulary scores were entered as additional fixed effects after entering the core model and trial number. Both LexTALE and VLT measured the vocabulary score of L2 speakers, and these two scores were highly correlated ($r(42) = 0.72, p < .001$). LexTALE and VLT scores were separately incorporated into models corresponding to each reading measurement. Given that these analyses showed very similar patterns and significance levels, we only reported models with the LexTale score as they provided a better fit for the data, based on BIC and AIC values. For the combined L1 and L2 speakers’ data, the full model included a core model, trial number, language-background factor (sum coded: L1 speakers: -1 versus L2 speakers: 1) and all interactions with language-background. All full models were subsequently reduced with the `drop1()` function, which performs model comparisons with a chi-square test. Note however that the two manipulated factors (phrase-type and word-order) were always kept in the reduced models to reflect the experimental design, even if their effects were not significant. The interactions with phrase-type,

word-order, language-background, vocabulary score and trial number were removed if they were not significant. Based on these tests, reduced models were reached to fit the data best. In the reduced models, variables with a t/z value above 1.96 are considered significant. The significance of interactions was assessed by model comparisons with the `anova` function. For all models, the `vif` function from the `car` package (Fox et al., 2023) was used to test for collinearity. All variance inflation factors were around 1, indicating a very low degree of correlation among variables.

For models including continuous dependent variables, when there was a main effect of collocation, we used Bayes factors (Kass & Raftery, 1995) to compare the relative evidence for models with and without interactions between phrase-type, word-order and language-background. The Bayes factors analysis was performed using the `ImBF` function from the `BayesFactor` package (Morey et al., 2022). When Bayes Factors (BFs) > 10, there is strong evidence for models including an interaction, while BFs < 1 provide evidence for models without an interaction. BFs between 3 and 10 provide weak to moderate evidence for models with an interaction. When there was a significant interaction, the `emmeans` function from the `emmeans` package (Lenth et al., 2024) was used to conduct pairwise comparisons of different levels. Lastly, we calculated Cohen’s d effect sizes using a method appropriate only for linear mixed effect models (Brysbaert & Stevens, 2018; Judd et al., 2017), which was based on estimated marginal means and total variance from the covariance model estimates.

3. Results

Reading time averages are shown in [Tables 1.A](#) and [1.B](#) for L1 and L2 speakers respectively. L1 speakers’ statistical results are reported in [Tables 2](#) and [3](#). L2 speakers’ statistical results are reported in [Tables 4](#) and [5](#). The statistical results from the combined L1 and L2 models are displayed in [Supplementary materials Table S3](#) and [Table S4](#).

Table 1.A. Summary of mean reading measures (ms) and total fixation counts in L1 speakers

Measures	L1 Speakers			
	Canonical constructions		Reversed constructions	
	Canonical collocations	Canonical novel word combinations	Reversed collocations	Reversed novel word combinations
Whole phrase				
First run reading time	632 (307)	650 (305)	567 (284)	577 (282)
Regression path reading time	719 (358)	736 (335)	688 (317)	704 (318)
Selective regression path reading time	689 (334)	714 (313)	648 (273)	666 (276)
Total reading time	869 (448)	931 (455)	854 (446)	930 (457)
Total fixation counts	4.12 (1.99)	4.45 (2.10)	4.11 (2.05)	4.45 (2.07)
Final word				
First fixation duration	217 (62)	224 (67)	221(71)	222 (67)
First run reading time	243 (92)	265 (113)	255 (105)	263 (108)
Regression path reading time	314 (182)	357 (213)	324 (199)	327 (182)
Selective regression path reading time	276 (119)	295 (126)	283 (124)	288 (126)
Total reading time	355 (192)	382 (218)	383 (226)	400 (228)
Total fixation counts	1.67 (0.88)	1.75 (0.92)	1.80 (1.04)	1.88 (1.05)

Note. The Standard Error of the Mean is shown in parentheses.

Table 1.B. Summary of mean reading measures (ms) and total fixation counts in L2 speakers

Measures	L2 Speakers			
	Canonical constructions		Reversed constructions	
	Collocations	Novel word combinations	Collocations	Novel word combinations
Whole phrase				
First run reading time	1093 (543)	1059 (569)	1046 (499)	1056 (539)
Regression path reading time	1272 (628)	1256 (575)	1256 (521)	1282 (588)
Selective regression path reading time	1218 (535)	1227 (550)	1211 (479)	1232 (554)
Total reading time	1797 (1201)	1945 (1186)	1890 (1249)	2025 (1307)
Total fixation counts	6.93 (4.28)	7.57 (4.34)	7.32 (4.38)	7.83 (4.65)
Final word				
First fixation duration	286 (98)	292 (102)	279 (98)	279 (99)
First run reading time	388 (158)	404 (170)	418 (200)	427 (202)
Regression path reading time	553 (358)	575 (353)	518 (288)	539 (351)
Selective regression path reading time	457 (192)	473 (196)	475 (236)	493 (250)
Total reading time	714 (507)	767 (500)	767 (553)	867 (598)
Total fixation counts	2.67 (1.80)	2.91 (1.86)	2.85 (1.88)	3.25 (2.09)

Note. The Standard Error of the Mean is shown in parentheses.

Table 2. L1 speakers: summary statistics for the phrase region

Factor	Statistics	FRRT	RPRT	SRPRT	TRT	TFC
Intercept [global mean]	β	608.51	714.22	681.77	899.94	1.42
	SE	23.57	33.62	30.83	40.83	0.04
	t/z	25.82	21.24	22.12	22.04	35.33
Phrase-type [collocation]	β	-5.47	-8.12	-9.70	-34.93	-0.04
	SE	7.85	8.06	8.14	13.59	0.01
	t/z	-0.69	-1.01	-1.19	-2.57*	-2.82*
	d	-0.04	-0.05	-0.06	-0.15	-
Word-order [canonical]	β	35.47	14.44	22.04	4.5	0.002
	SE	9.32	8.05	8.64	12.73	0.01
	t/z	3.8*	1.79	2.55*	0.36	0.12
	d	0.24	0.09	0.14	0.02	-

Note. FRRT: first run reading time, RPRT: regression path reading time, SRPRT: selective regression path reading time, TRT: total reading time, TFC: total fixation count.
* = $p < .05$.

3.1. L1 speakers results

Phrase analysis. Table 2 shows that the phrase-type effect was observed on total readings times ($\beta = -34.93$, $d = -0.15$, $t = -2.57$, $p = .011$) and total fixation counts ($\beta = -0.04$, $z = -2.82$, $p = .004$): L1 speakers read collocations faster and fixated fewer times on collocations than novel word combinations. The main effect of word-order was found on first run reading times ($\beta = 35.47$, $d = 0.24$, $t = 3.8$, $p < .001$) and selective regression times ($\beta = 22.04$, $d = 0.14$, $t = 2.55$, $p = .013$), with reversed constructions being processed faster than canonical ones. No significant main effects or interactions were observed for the other measures (all $ps > .7$). Furthermore, there was no interaction between phrase-type and word-order on total readings times ($\chi^2(1) = 0.15$, $p = .697$; BFs = 0.09) and total fixation counts ($\chi^2(1) = 0.072$, $p = .788$). The significant main effect of phrase-type and

the lack of an interaction with word-order indicates that L1 speakers consistently read collocations faster than novel word combinations, regardless of their word order.

Final word analysis. For L1 speakers (summarised in Table 3), the phrase-type effect in the canonical constructions was significant in the first run time model ($\beta = -10.22$, $d = -0.2$, $t = -2.4$, $p = .021$) and in the regression path reading time model ($\beta = -20.11$, $d = -0.2$, $t = -2.45$, $p = .017$). The remaining eye-tracking measures did not show any significant effects.

3.2. L2 speakers results

Phrase analysis. There was a phrase-type effect on total reading times ($\beta = -71.57$, $d = -0.12$, $t = -2.67$, $p = .008$) and total fixation counts ($\beta = -0.04$, $z = -2.92$, $p = .003$), with collocations being

Table 3. L1 speakers: summary statistics for the final word region

Factor	Statistics	FFRT	FRRT	RPRT	SRPRT	TRT	TFC
Final words: canonical constructions [nouns]							
Intercept [global mean]	β	219.57	252.25	332.72	283.14	368.3	0.53
	SE	4.37	6.97	15.07	9.28	16.06	0.03
	t/z	50.31	36.2	22.07	30.51	22.93	15.09
Phrase-type [collocation]	β	-3.18	-10.22	-20.11	-8.54	-13.78	-0.02
	SE	2.64	4.27	8.19	5.12	8.25	0.03
	t/z	-1.21	-2.4*	-2.45*	-1.67	-1.67	-0.79
	d	-0.1	-0.20	-0.2	-0.14	-0.13	-
Final words: reversed constructions [verbs]							
Intercept [global mean]	β	221.20	258.48	324.84	285.33	389.7	0.58
	SE	5.14	7.49	12.50	9.36	18.34	0.04
	t/z	42.99	34.54	25.95	30.49	21.24	14.3
Phrase-type [collocation]	β	-0.60	-3.87	-1.87	-2.81	-9.85	-0.02
	SE	2.57	3.72	7.15	4.65	9.26	0.03
	t/z	-0.27	-1.04	-0.26	-0.61	-1.06	-0.80
	d	-0.02	-0.07	-0.02	-0.04	-0.09	-

Note. FFRT: first fixation reading time, FRRT: first run reading time, RPRT: regression path reading time, SRPRT: selective regression path reading time, TRT: total reading time, TFC: total fixation count. * = $p < .05$.

Table 4. L2 speakers: summary statistics for the phrase region

Factor	Statistics	FRRT	RPRT	SRPRT	TRT	TFC
Intercept [global mean]	β	1065.99	1270.64	1225.39	1919.74	1.92
	SE	39.72	53.94	51.29	127.16	0.06
	t/z	26.84	23.55	23.89	15.1	32.38
Phrase-type [collocation]	β	4.544	-5.73	-10.8	-71.57	-0.04
	SE	13.86	14.71	12.95	26.79	0.01
	t/z	0.33	-0.39	-0.78	-2.67*	-2.92*
	d	0.02	-0.02	-0.04	-0.12	-
Word-order [canonical]	β	11.54	-4.29	-0.14	-42.94	-0.01
	SE	15.29	15.92	13.63	31.32	0.02
	t/z	0.75	-0.27	-0.01	-1.4	-0.92
	d	0.04	-0.02	-0.001	-0.07	-
Vocabulary-score	β	-134.31	-185.80	-172.31	-344.08	-0.17
	SE	38.48	50.80	50.49	120.05	0.06
	t/z	-3.51*	-3.66*	-3.4*	-2.87*	-2.90*
	d	-0.27	-0.33	-0.34	-0.29	-

Note. FRRT: first run reading time, RPRT: regression path reading time, SRPRT: selective regression path reading time, TRT: total reading time, TFC: total fixation count. * = $p < .05$.

processed more quickly than novel word combinations (see Table 4). The main effect of word-order was not significant on any measure. There was no interaction between phrase-type and word-order on total reading times ($\chi^2(1) = 0.01, p = .996; BFs = 0.006$) and total fixation counts ($\chi^2(1) = 0.35, p = .549$). This means that the collocational processing advantage was independent from word order in L2 speakers as well. No other effects approached significance. There was an effect of vocabulary score, which was measured by LexTALE, on all eye-

tracking measures. The higher the score participants had, the faster they read. However, we did not detect any interactions between the vocabulary score and phrase-type on any measures.

Final word analysis. Table 5 shows that there was a phrase-type effect in the total reading time for reversed constructions ($\beta = -48.28, t = -2.07, p = .04$) and in the total fixation count measures for canonical constructions ($\beta = -0.05, z = -2.25, p = .024$) and reversed constructions ($\beta = -0.06, z = -2.45, p = .014$). Participants had fewer

Table 5. L2 speakers: summary statistics for the final word region

Factor	Statistics	FFRT	FRRT	RPRT	SRPRT	TRT	TFC
Final words: canonical constructions [nouns]							
Intercept [global mean]	β	287.57	393.11	558.09	461.38	736.57	0.96
	SE	5.97	10.77	26.32	15.07	46.63	0.05
	t/z	48.2	36.49	21.2	30.62	15.8	18.33
Phrase-type [collocation]	β	2.73	-9.32	-10.95	-8.31	-24.58	-0.05
	SE	3.51	6.1	13.88	7.92	20.76	0.02
	t/z	0.78	-1.53	-0.79	-1.05	-1.18	-2.25*
	d	0.05	-0.12	-0.06	-0.09	-0.1	-
Vocabulary-score	β	-18.83	-52.52	-84.67	-61.23	-129.89	-0.16
	SE	5.89	10.3	25.1	14.13	44.97	0.05
	t/z	-3.2*	-5.01*	-3.37*	-4.33*	-2.89*	-3.07*
	d	-0.19	-0.34	-0.24	-0.33	-0.26	-
Vocabulary-score* Phrase-type	β						-0.05
	SE						0.02
	t/z						-2.23*
Final words: reversed constructions [verbs]							
Intercept [global mean]	β	278.4	423.89	529.36	484.11	813.69	1.03
	SE	6.21	15.13	20.68	17.26	56.32	0.06
	t/z	44.8	28.01	25.59	28.04	14.45	17.45
Phrase-type [collocation]	β	0.04	-4.64	-9.59	-9.19	-48.28	-0.06
	SE	3.98	8.89	13.46	9.77	23.24	0.03
	t/z	0.01	-0.52	-0.72	-0.94	-2.07*	-2.45*
	d	0.001	-0.05	-0.06	-0.08	-0.17	-
Vocabulary-score	β	-17.7	-47.57	-76.65	-63.45	-174.41	-0.16
	SE	6.07	13.86	18.69	16.05	53.67	0.06
	t/z	-2.92*	-3.43*	-4.1*	-3.95*	-3.25*	-2.88*
	d	-0.18	-0.23	-0.25	-0.27	-0.32	-

Note. FFRT: first fixation reading time, FRRT: first run reading time, RPRT: regression path reading time, SRPRT: selective regression path reading time, TRT: total reading time, TFC: total fixation count.

*= $p < .05$.

fixations on the final word in collocations than in novel word combinations, and when fixated, they spent less time reading the collocations. The interaction between vocabulary score and phrase-type was significant for the canonical constructions for the total fixation count measure ($\beta = -0.05$, $z = -2.23$, $p = .026$). L2 speakers with higher vocabulary scores were more likely to show fewer fixations on the final noun of collocations, relative to novel word combinations, while participants with lower vocabulary scores were more likely to show the opposite pattern (Figure 2).

3.3. Comparison of L1 and L2 speakers

After looking at L1 and L2 models separately, we built models with L1 and L2 data to investigate whether there are differences in collocation processing between the two groups other than the expected longer reading times for L2 speakers. Figure 3 shows the comparison between L1 and L2 speakers in reading collocations and novel word combinations on the phrase-level total reading time measure.

Phrase analysis. The L1 and L2 data were combined into one model. The L1 and L2 speakers differed in all measures (see Supplementary materials Table S3). As expected, L2 speakers took longer to read than L1 speakers. The phrase-type effect was observed in total reading times ($\beta = -53.62$, $d = -0.11$, $t = -3.20$, $p = .001$) and total fixation counts ($\beta = -0.04$, $z = -3.42$, $p < .001$), with collocations being read faster than novel word combinations. There was also an effect of word-order on first run reading times ($\beta = 24.11$, $d = 0.11$, $t = 2.43$, $p = .016$), with reversed constructions processed faster than canonical ones. The interaction between language-background and phrase type was not significant on total reading times ($\chi^2(1) = 2.80$, $p = .089$; BFs = 0.12) or fixation count ($\chi^2(1) = 0.002$, $p = .958$). This indicates that L1 and L2 participants read collocations qualitatively similarly. There was no three-way interaction between language-background, word-order and phrase-type on any measures.

Final word analysis. The language-background variable was a consistent predictor on all measures (see Supplementary materials

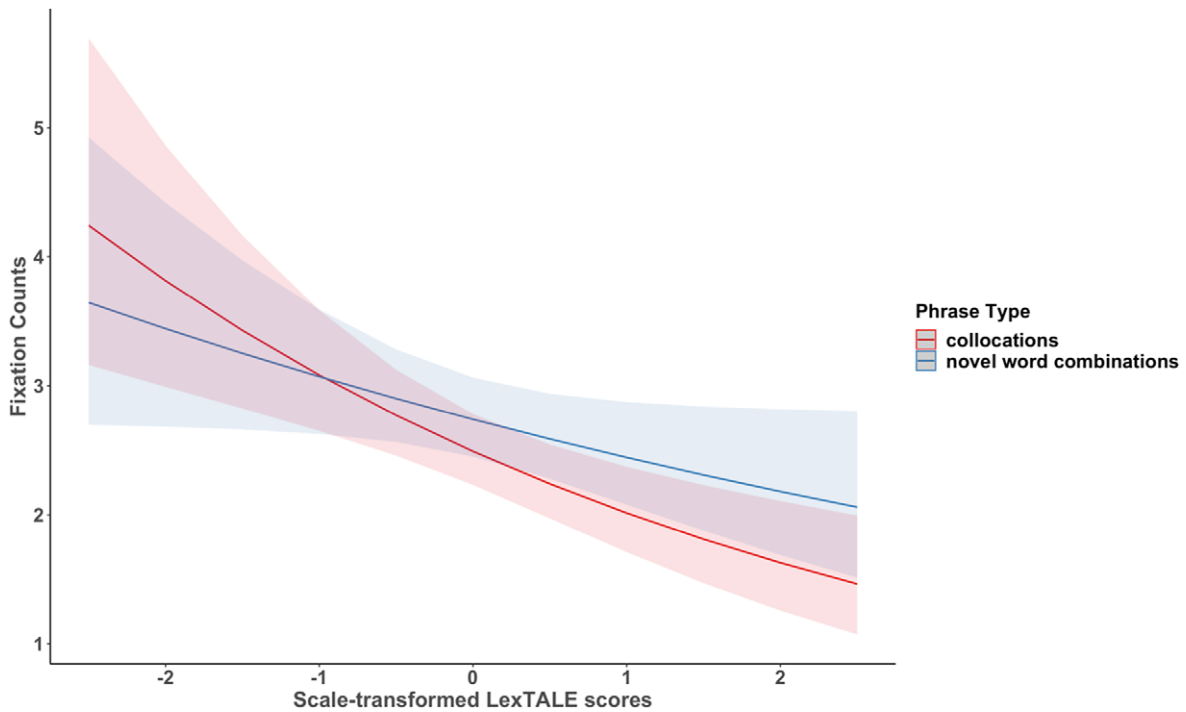


Figure 2. L2 speakers: interaction between vocabulary score and phrase-type in total fixation count for canonical constructions.

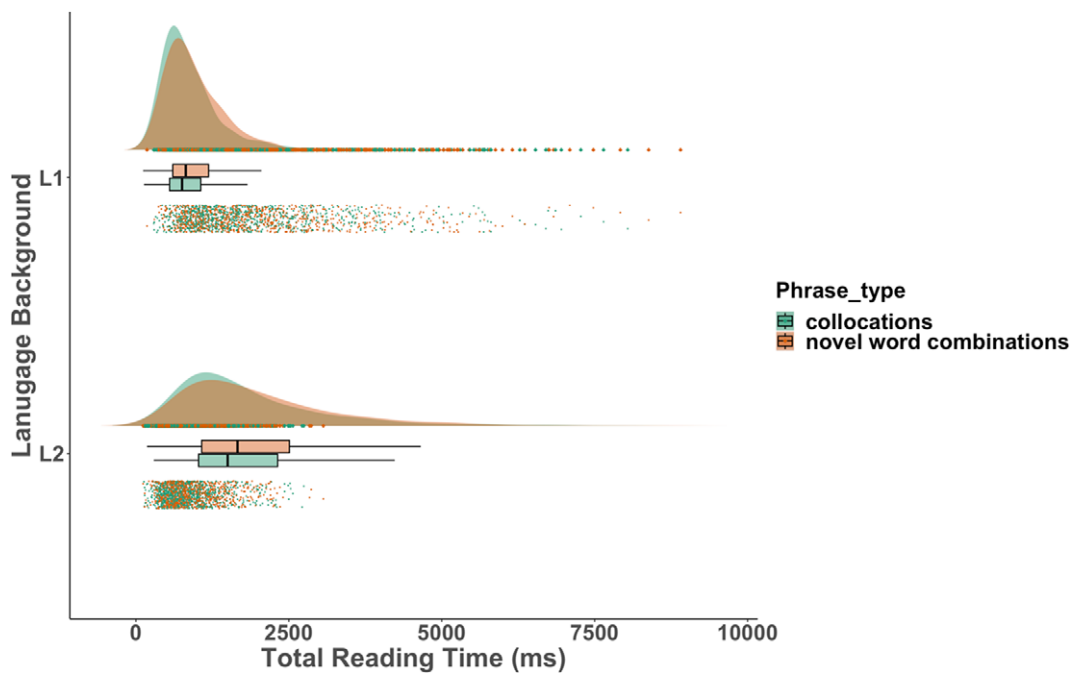


Figure 3. Total phrase reading time for collocations and novel word combinations by language background.

Table S4), with L2 speakers reading final words more slowly, and fixating them more often than L1 speakers in both canonical and reversed constructions. The phrase-type effect, with collocations being read faster than novel word combinations, was observed in first run reading times for canonical constructions ($\beta = -9.95$, $d = -0.15$, $t = -2.37$, $p = .02$) and in the total reading time and fixation count ($\beta = -29.54$, $t = -2.05$, $p = .048$; $\beta = -0.04$, $z = -2.02$,

$p = .043$) for reversed constructions. There was only an interaction between phrase-type and language-background on the total reading times for reversed constructions ($\chi^2(1) = 5.64$, $p = .017$), where a phrase-type effect found in L2 speakers ($\beta = -97.4$, $t = -2.94$, $p = .004$) but not L1 speakers ($\beta = -21.7$, $t = -0.65$, $p = .51$). We also did not observe a three-way interaction between language-background, phrase-type, and word-order.

4. Discussion

The current study tested the effects of changing word order on collocation processing in L1 and L2 speakers while controlling for contextual predictability. Results show that collocations were processed faster than novel word combinations, confirming previous studies of collocation advantage using the eye-tracking-while-reading paradigm (Li et al., 2021; Sonbul, 2015; Vilkaitė, 2016; Vilkaitė & Schmitt, 2019) and using behavioural tasks (e. g., Öksüz et al., 2021; Wolter & Gyllstad, 2011, 2013; Wolter & Yamashita, 2015, 2018) and extend these findings to contexts where predictability was controlled for. While the collocation processing advantage mainly appeared in later reading measures, L1 speakers also showed the advantage from canonical collocations on the first run and regression path reading. Crucially, there was no interaction between phrase type and word order on phrase-level processing, suggesting that reversed collocations (in non-canonical word order) still showed an advantage. This study also provides initial evidence that advanced L2 speakers process both canonical and reversed collocations similarly to L1 speakers.

4.1. Collocation facilitation effect

The study confirmed that collocations are processed faster than novel word combinations (Arnon & Snider, 2010; Carrol & Conklin, 2014b, 2020; Li et al., 2021; Öksüz et al., 2021; Siyanova & Schmitt, 2008; Sonbul, 2015; Vilkaitė, 2016; Vilkaitė & Schmitt, 2019; Wolter & Gyllstad, 2011, 2013). Because cloze probability was controlled for, this collocation effect is independent from contextual predictability. Results show that the collocation effect exists in both early (first-run reading times and selective regression reading times, for L1 speakers only) and late reading measures (total reading times and total fixation counts for both L1 and L2 speakers), though late eye-tracking measures demonstrated a more consistent effect. This generally aligns with previous evidence in the processing of MWS, indicating that language users are sensitive not only to word-level but also phrase-level frequency, such as lexical bundles, idioms, and collocations (Arnon et al., 2017; Arnon & Snider, 2010). Previous studies found the facilitation effect from reading verb-noun collocations with base verbs (McDonald & Shillcock, 2003a, 2003b; Vilkaitė, 2016). The present study found that collocations can also elicit faster processing in their past tense. This fits in with (Vilkaitė-Lozdienė 2022) finding that cumulative exposure to a collocation in all different morphological forms contributes to its faster processing. Thus, the finding supports the usage-based approach to L1 acquisition (Bybee, 1998; Goldberg, 1995; Tomasello, 2003, 2009) and L2 acquisition (Ellis & Wulff, 2020), which proposes that language users acquire a set of linguistic constructions of varying form, size, and complexity and that all these constructions are similarly affected by frequency.

Phrase-level and final word analyses provided different evidence. In the phrase-level analyses, there was a consistent advantage for collocations over novel word combinations on late eye-tracking measures for both L1 and L2 speakers. However, final word analyses showed a different pattern: canonical collocations affected earlier measures (first-run reading time and regression path reading time) only for L1 speakers. Overall, these findings are consistent with more recent eye-tracking studies (Li et al., 2021; Sonbul, 2015; Vilkaitė, 2016; Vilkaitė-Lozdienė, 2022) and suggest that collocational status can affect an early stage of phrase processing. However, the present study generally observed a less robust and less consistent collocation effect on early measures compared to

previous studies, probably because contextual predictability was kept minimal and matched.

This study also sheds light on previous findings by Frisson et al. (2005) who found that the transitional probability effect no longer influenced the early stages of verb-noun pairs processing once contextual predictability was controlled. However, it is worth noting that overall the verb-noun pairs in their study had a much lower average phrase-level frequency compared to recent collocation studies. This suggests that the association between verbs and nouns in their study was weaker than in collocation studies, which might have led to weak activation of the second words when encountering the first words of those pairs. Exposure to highly frequently co-occurring words results in deeper memory entrenchment, making these co-occurring words become more “lexicalised units” and “templated” in the mental lexicon. It is possible that verb-noun pairs with lower phrase-level frequency are less entrenched in memory, so that, when encountering their collocational form, it takes longer to match it to their collocational meaning. This also suggests that an early collocation effect might be restricted to high-frequency combinations only, and future research should establish whether this is a monotonous effect, with an increasingly stronger collocation effect as the frequency of the combination increases, or whether there is cut-off where an MWS becomes a “true” collocation. To test this, it is essential to control the contextual predictability of all parts of a collocation.

4.2. Facilitation effect in reversed collocations

The study revealed that collocations maintain a processing advantage when word order changes. There were consistent results in the phrase-level processing in L1 and L2 speakers: phrase type was a robust predictor and there was no interaction between phrase type and word order. This suggests that collocations are processed faster than novel word combinations regardless of word order variation.

However, final word analyses indicated that while the collocation advantage was present in the early measures of processing in L1 speakers, the advantage in reversed collocations was limited to the later measures. This suggests that canonical collocations have an earlier effect than reversed ones. Because reversed collocations are much less frequent than canonical constructions (*he brought peace* versus *peace was brought*), this suggests that while the general collocation effect is meaning-driven, canonical forms have an added frequency-related advantage. This finding of faster processing of reversed collocations than their controls contrasts with the absence of backward priming in English collocations reported by (Vilkaitė-Lozdienė and Conklin 2021), who found that word [n] (e.g., *attract*) primes word [n + 1] (*attention*), while word [n + 1] does not facilitate word [n]. However, as the authors explained, this absence may be due to reversing verb-object phrases (e.g., *attention attract*) creating unnatural or implausible phrases and grammatical errors.

As noted in the introduction, two previous studies (Vilkaitė, 2016; Vilkaitė-Lozdienė, 2022) demonstrated that collocations are flexible and can be modified by inserting words or changing morphological forms without disrupting online processing in L1 speakers. The present study extended this finding to reversed word order, showing that it does not reduce the collocation facilitation effect at the phrase processing level compared to controls. This study also extended (Kyriacou et al. 2020) findings from idioms to collocations and from L1 to L2 speakers. Kyriacou et al. found that L1 speakers read reversed idioms faster and with fewer regressions than reversed controls, but were still slower with passive than active

idioms, suggesting that idioms' faster reading is due to activating their figurative meaning, even in non-canonical forms. In contrast, the present study found that reversed collocations were read faster than their canonical counterparts. This faster processing of passive than active collocations is perhaps less surprising than it may appear. First, several studies found no consistent processing advantage for active over passive sentences (Carrithers, 1989; Paolazzi et al., 2019, 2022; Traxler et al., 2014), and Paolazzi et al. (2019) reported consistently shorter reading times for passive than active sentences in four self-paced reading experiments. Second, the present study only included agent-patient verbs; Paolazzi et al. (2022) noted that processing difficulty in passive sentences varies with verb type, and passivized subject-experiencer verbs (stative predicates, e.g., *love*, *admire*) result in longer fixation durations but passivized agent-patient verbs do not. Additionally, unlike Kyriacou et al. (2020), the preceding contexts of the experimental sentences in the current study provided little information, reflected in the low predictability scores across conditions. While predictability was low overall, other variables could not be controlled across active/passive constructions. For example, inserting a subject pronoun in the active constructions with little preceding contextual information (e.g., *they* in *Margaret is not sure that they brought peace/meat to the village*) might have caused participants to spend extra time resolving its reference, slowing down reading in the active voice conditions. Together with Kyriacou et al. (2020), the present findings of a collocation effect for reversed word order support the idea that formulaic language constructions are open to variation and modification.

4.3. The question of holistic storage

This finding suggests that the processing advantage in reversed collocations is most likely due to the activation of the collocation's meaning rather than form. Examining evidence against the activation of collocational forms, our finding challenges holistic storage, where form and meaning are stored as a whole (e.g., Goldinger, 1996; Jiang & Nekrasova, 2007; Wray, 2002). First, if a collocation's form (e.g., *he brought peace*) is stored in the mental lexicon and retrieved as a whole chunk, changing word order (e.g., *peace was brought*) should disrupt processing. On the contrary, the present study found that reversed collocations were processed faster. Secondly, one might argue that reversed collocations have their own lexical entries in memory, which could facilitate processing. Yet, the reversed collocations in our study were rare (all MI scores of collocations in different morphological forms were lower than 1.5). Moreover, Vilkaitė-Lozdienė (2022) used infrequent collocations in participle form that are unlikely to be stored as holistic units, yet the collocation advantage was still observed. Combining the present study with Vilkaitė-Lozdienė's finding, we argue that the collocational form is probably not stored as a complete unit.

Looking at evidence supporting the role of collocational meaning, the observed processing advantage for reversed collocations suggests that the unitary semantic representation contributes to their faster processing. (Vilkaitė 2016) finding may be taken as counterevidence, as the researcher found a processing advantage in non-adjacent collocations, where inserted words might disrupt the collocational meaning. Yet, most inserted words only modified the nouns of the collocations (e.g., *the degree of*, *at least some*, *a bit of*, *some of the*, *the best possible*), adding information without changing the collocation's overall meaning. Thus, (Vilkaitė 2016) findings cannot be taken to refute the central role of collocational meaning.

Finally, we must acknowledge the complexity of separating form and meaning activation in collocation processing. While our findings indicate meaning activation as the primary factor, collocations likely involve interconnected representations of their components (*bring* and *peace*) and the entire collocation (*bring peace*). These representations are assumed to be bidirectionally connected, allowing activation to spread between the components and the chunk (Vilkaitė-Lozdienė & Conklin, 2021). The processing advantage may stem from a dual-route mechanism, where both the components and the chunk are activated simultaneously. The pathway leading to faster processing – whether through the components or the chunk – determines the outcome. This also applies to reversed forms (e.g., *peace was brought*), assuming that individual components and the chunk are activated, even when their order is changed.

4.4. Similarities between L1 and L2 speakers

This study showed that changing the word order of a collocation has similar effects on L1 and advanced L2 speakers. L2 speakers showed a processing advantage for collocations in later processing measures at both phrase and word levels. This aligns with previous finding that proficient L2 speakers are sensitive to phrase-level frequency or association strength between words (e.g., Öksüz et al., 2021; Sonbul, 2015; Vilkaitė & Schmitt, 2019). Aligning with the discussion above, the lack of an early effect on the final word may relate to differences in collocational frequency between L1 and L2 speakers. It is safe to assume that L2 speakers receive less exposure to the collocations of the target language than L1 speakers. If indeed the early effect is affected by frequency, it is not surprising that this effect is stronger in L1 than L2 speakers.

Importantly, the phrase-level facilitation effect was observed for both canonical and reversed collocations. This provides initial evidence that experienced L2 speakers process variation collocations in a qualitatively similar manner to L1 speakers, even though L2 speakers develop collocation competence very slowly (Groom, 2009; Laufer & Waldman, 2011). This also questions the claim that L1 and L2 speakers process MWs, such as collocations, fundamentally differently from L1 speakers, with L1 speakers assigning meaning to larger linguistic chunks and L2 speakers heavily relying on words and rules (Wray, 2002, 2008). Instead, our findings suggest that L1 speakers and proficient L2 speakers process large linguistic constructions in a similar manner, activating the unitary semantic meaning.

These results contrast with (Vilkaitė and Schmitt 2019) finding that the processing advantage in non-adjacent collocations disappeared in L2 speakers. The divergent findings cannot be solely attributed to differences in proficiency, as the L2 speakers in our study had a lower average Vocabulary Levels Test score than the participants in their study. Rather, this processing difference might reflect the difference in the constructions used. It is likely that the L2 speakers in Vilkaitė and Schmitt's study initially activated the second word of the collocation after encountering the first word, but the intervening words between the first and second word complicated the integration of semantic information, weakening the activation of the collocational meaning. Overall, this shows that the facilitation effect of variation collocations depends on the type of collocation modification and word proximity.

4.5. The proficiency effect on L2 collocation processing

We found an interaction in later measures on final word processing between vocabulary score and phrase type. More proficient L2

speakers fixated on the final word of collocations for a shorter time than on the final word of novel word combinations, while less proficient ones (those with lower vocabulary levels) spent more time fixating the final word of collocations. This interaction suggests that unitary semantic activation is moderated by L2 proficiency. On the one hand, less proficient L2 speakers might prefer to analyse individual components and then compute the meaning online. However, we do not deny the possibility of delayed unitary activation of meaning in less proficient L2 speakers. Greater dependence on analysing and computing individual collocation components can slow down processing. On the other hand, proficient L2 speakers have more exposure to collocations, likely resulting in frequently encountered collocations becoming lexicalised and forming their own unitary semantic representations in the mental lexicon. Thus, more proficient L2 speakers can either compute the familiar collocational meaning more quickly or activate its unitary semantic representation more straightforwardly. However, distinguishing between these two explanations is difficult in the current study.

In addition, our study also observed a main effect of vocabulary knowledge in all L2 statistical models, with L2 speakers with better vocabulary knowledge processing both collocations and novel word combinations faster than those with smaller vocabularies. This main effect of vocabulary level is expected since vocabulary size is closely related to general L2 proficiency, at least in English (Alderson & Huhta, 2005; Milton, 2013), and more proficient speakers tend to read faster (Berzak et al., 2018; Godfroid et al., 2020; Muñoz, 2017).

In conclusion, the present findings add to the growing body of literature that supports the collocation advantage in both L1 and L2 speakers. Importantly, changing collocations' word order, namely from canonical to reversed passivized construction, did not disrupt collocation processing in either L1 or L2 speakers. This suggests that both L1 and proficient L2 speakers can maintain the collocation processing advantages with variation collocations. While these findings are promising, future studies could address some limitations. First, only English verb-noun collocations were tested, and it remains unknown whether languages with stricter or looser word order show similar trends (see e.g., (Vilkaitė-Lozdienė and Conklin 2021)). Second, items in this study were a combination of verb noun pairs varying by noun (e.g., *bring peace/meat*) and by verb (e.g., *create/produce condition*). Although statistical models found no differences between the two types of verb-noun pairs, future studies could improve the experimental design by ensuring consistent variation. Third, whether reversing other types of collocations (e.g., adjective-nouns such as *fatal mistakes – mistakes that are fatal*) retains an advantage is unclear and might depend on the semantic restrictedness of the expression. Fourth, testing L2 speakers with same or different linguistic and orthographic backgrounds could reveal whether the magnitude of semantic activation remains consistent when processing reversed collocations. Finally, including L2 users with varying proficiency levels and different L1 backgrounds would help generalise our results.

Supplementary material. The supplementary material for this article can be found at <http://doi.org/10.1017/S1366728925000057>.

Data availability. The data that support the findings of this study are openly available in [reversed L1 and L2 collocations] at <https://osf.io/5br6v/>

Acknowledgement. We would like to express our gratitude to Dr Mahmoud Elsherif for his help in checking R codes, analysis and reading the first draft. We

also thank the anonymous reviewers for their detailed feedback, and all the participants of the study.

Competing interest. The authors declare no competing interests.

References

- Alderson, J. C., & Huhta, A. (2005). The development of a suite of computer-based diagnostic tests based on the Common European Framework. *Language Testing*, 22(3), 301–320. <https://doi.org/10.1191/0265532205lt310oa>
- 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 multiword 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.
- 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>
- Berzak, Y., Katz, B., & Levy, R. (2018). *Assessing language proficiency from eye movements in reading* (arXiv:1804.07329). arXiv. <https://doi.org/10.48550/arXiv.1804.07329>
- Biber, D., Johansson, S., Leech, G., & Conrad, S. (1999). *Longman grammar of spoken and written English*. London: Longman.
- Brybaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition*, 1(1). <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.), *The panels: Vol. CLS 34* (Issue 2). University of Chicago: Chicago Linguistic Society.
- Carrol, G., & Conklin, K. (2014a). Eye-tracking multi-word units: Some methodological questions. *Journal of Eye Movement Research*, 7(5). <https://doi.org/10.16910/jemr.7.5.5>
- Carrol, G., & Conklin, K. (2014b). 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. (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>
- Clifton, C., Staub, A., & Rayner, K. (2007). Eye movements in reading words and sentences. In R. P. G. van Gompel, W. S. Murray, & R. L. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 341–371). Elsevier.
- Cowie, A. P. (1998). *Phraseology: Theory, analysis, and applications*. Oxford University Press.
- Durrant, P., & Doherty, A. (2010). Are high-frequency collocations psychologically real? Investigating the thesis of collocational priming. *Corpus Linguistics and Linguistic Theory*, 6(2), 125–155. <https://doi.org/10.1515/clt.2010.006>
- Ellis, N. C., & Wulff, S. (2020). Usage-based approaches to L2 acquisition: An introduction. In B. VanPatten, G. Keating, & S. Wulff (Eds.), *Theories in second language acquisition* (pp. 63–82). Routledge. <https://doi.org/10.4324/9780429503986>
- Fox, J., Weisberg, S., Price, B., Adler, D., Bates, D., Baud-Bovy, G., Bolker, B., Ellison, S., Firth, D., Friendly, M., Gorjanc, G., Graves, S., Heiberger, R., Krivitsky, P., Laboissiere, R., Maechler, M., Monette, G., Murdoch, D., Nilsson, H., ... R-Core. (2023). *car: Companion to Applied Regression* (Version 3.1–2) [Computer software]. <https://CRAN.R-project.org/package=car>
- Frisson, S., Rayner, K., & Pickering, M. J. (2005). Effects of contextual predictability and transitional probability on eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 862–877. <https://doi.org/10.1037/0278-7393.31.5.862>
- Gablasova, D., Brezina, V., & McEnery, T. (2017). Collocations in corpus-based language learning research: Identifying, comparing, and interpreting the evidence. *Language Learning*, 67(S1), 155–179. <https://doi.org/10.1111/lang.12225>

- Godfroid, A., Winke, P., & Conklin, K. (2020). Exploring the depths of second language processing with eye tracking: An introduction. *Second Language Research*, 36(3), 243–255. <https://doi.org/10.1177/0267658320922>
- Goldberg, A. (1995). *Constructions: A construction grammar approach to argument structure*. University of Chicago Press.
- Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 1166–1183. <https://doi.org/10.1037/0278-7393.22.5.1166>
- Groom, N. (2009). Effects of Second Language Immersion on Second Language Collocational Development. In A. Barfield & H. Gyllstad (Eds.), *Researching Collocations in Another Language: Multiple Interpretations* (pp. 21–33). Palgrave Macmillan UK. https://doi.org/10.1057/9780230245327_2
- Gyllstad, H., & Wolter, B. (2016). Collocational processing in light of the phraseological continuum model: Does semantic transparency matter? *Language Learning*, 66(2), 296–323. <https://doi.org/10.1111/lang.12143>
- Holcomb, P. J. (1988). Automatic and attentional processing: An event-related brain potential analysis of semantic priming. *Brain and Language*, 35(1), 66–85. [https://doi.org/10.1016/0093-934X\(88\)90101-0](https://doi.org/10.1016/0093-934X(88)90101-0)
- Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford University Press.
- Howarth, P. (1998). Phraseology and second language proficiency. *Applied Linguistics*, 19(1), 24–44. <https://doi.org/10.1093/applin/19.1.24>
- Hunston, S. (2022). *Corpora in applied linguistics* (2nd ed.). Cambridge University Press.
- Jiang, N., & Nekrasova, T. M. (2007). The processing of formulaic sequences by second language speakers. *The Modern Language Journal*, 91(3), 433–445. <https://doi.org/10.1111/j.1540-4781.2007.00589.x>
- Judd, C. M., Westfall, J., & Kenny, D. A. (2017). Experiments with More Than One Random Factor: Designs, Analytic Models, and Statistical Power. *Annual Review of Psychology*, 68(Volume 68, 2017), 601–625. <https://doi.org/10.1146/annurev-psych-122414-033702>
- Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90(430), 773–795. <https://doi.org/10.1080/01621459.1995.10476572>
- Kilgarriff, A., Rychly, P., Smrz, P., & Tugwell, D. (2004). “The Sketch Engine.” In G. Williams & S. Vessier (Eds.), *Proceedings of Euralex 2004* (pp. 105–116). Université de Bretagne-Sud.
- Kiss, G. R., Armstrong, C., Milroy, R., & Piper, J. (1973). An associative thesaurus of English and its computer analysis. *The Computer and Literary Studies*, 153.
- Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, 4(12), 463–470. [https://doi.org/10.1016/S1364-6613\(00\)01560-6](https://doi.org/10.1016/S1364-6613(00)01560-6)
- Kuznetsova, A., Brockhoff, P. B., Christensen, R. H. B., & Jensen, S. P. (2020). *Package 'lmerTest'*. (Version 3.1–3) [Computer software]. <https://CRAN.R-project.org/package=lmerTest>
- 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>
- Laufer, B., & Waldman, T. (2011). Verb-Noun collocations in second language writing: A Corpus analysis of learners. *English*. *Language Learning*, 61(2), 647–672. <https://doi.org/10.1111/j.1467-9922.2010.00621.x>
- Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods*, 44(2), 325–343. <https://doi.org/10.3758/s13428-011-0146-0>
- 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>
- 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>
- McDonald, S. A., & Shillcock, R. C. (2003a). Eye movements reveal the on-line computation of lexical probabilities during reading. *Psychological Science*, 14(6), 648–652. https://doi.org/10.1046/j.0956-7976.2003.psci_1480.x
- McDonald, S. A., & Shillcock, R. C. (2003b). Low-level predictive inference in reading: The influence of transitional probabilities on eye movements. *Vision Research*, 43(16), 1735–1751. [https://doi.org/10.1016/S0042-6989\(03\)00237-2](https://doi.org/10.1016/S0042-6989(03)00237-2)
- Milton, J. (2013). Measuring the contribution of vocabulary knowledge to proficiency in the four skills. In C. Bardel, C. Lindqvist, & B. Laufer (Eds.), *L2 vocabulary acquisition, knowledge and Use: New Perspectives on Assessment and Corpus Analysis* (pp. 57–78). Eurosla Monographs Series, 2.
- Molinaro, N., Canal, P., Vespignani, F., Pesciarelli, F., & Cacciari, C. (2013). Are complex function words processed as semantically empty strings? A reading time and ERP study of collocational complex prepositions. *Language and Cognitive Processes*, 28(6), 762–788. <https://doi.org/10.1080/01690965.2012.665465>
- Moon, R. (1998). *Fixed expressions and idioms in English: A corpus-based approach*. Clarendon Press.
- Morey, R. D., Rouder, J. N., Jamil, T., Urbanek, S., Forner, K., & Ly, A. (2022). *Package 'BayesFactor': Computation of Bayes Factors for Common Designs* (Version 0.9.12–4.4) [Computer software]. <https://CRAN.R-project.org/package=BayesFactor>
- Muñoz, C. (2017). The role of age and proficiency in subtitle reading. An eye-tracking study. *System*, 67, 77–86. <https://doi.org/10.1016/j.system.2017.04.015>
- Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, 24(2), 223–242. <https://doi.org/10.1093/applin/24.2.223>
- 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>
- Pinker, S. (1991). Rules of language. *Science*, 253(5019), 530–535. <https://doi.org/10.1126/science.1857983>
- Pinker, S., & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition*, 28(1), 73–193. [https://doi.org/10.1016/0010-0277\(88\)90032-7](https://doi.org/10.1016/0010-0277(88)90032-7)
- Powell, M. J. (2009). *The BOBYQA algorithm for bound constrained optimization without derivatives* (26; Cambridge NA Report NA2009/06). University of Cambridge.
- R Core Team. (2022). *R: A language and environment for statistical* (Version 4.2.1) [Computer software]. <https://www.R-project.org/>
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372–422. <https://doi.org/10.1037/0033-2909.124.3.372>
- Rayner, K., Kambe, G., & Duffy, S. A. (2000). The effect of clause wrap-up on eye movements during reading. *Quarterly Journal of Experimental Psychology*, 53A(4), 1061–1080. <https://doi.org/10.1080/713755934>
- 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.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford University Press, USA.
- 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>
- 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>
- Taylor, W. L. (1953). “Cloze procedure”: A new tool for measuring readability. *Journalism & Mass Communication Quarterly*, 30(4), 415–433. <https://doi.org/10.1177/107769905303000401>
- Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Harvard university press.
- Tomasello, M. (2009). The usage-based theory of language acquisition. In E. L. Bavin (Ed.), *The Cambridge handbook of child language* (pp. 69–87). Cambridge University Press. <https://doi.org/10.1017/CBO9780511576164.005>
- Veldre, A., & Andrews, S. (2014). Lexical Quality and Eye Movements: Individual Differences in the Perceptual Span of Skilled Adult Readers. *Quarterly Journal of Experimental Psychology*, 67(4), 703–727. <https://doi.org/10.1080/17470218.2013.826258>

- 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.** (2022). Do different morphological forms of collocations show comparable processing facilitation? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *48*, 1328–1347. <https://doi.org/10.1037/xlm0001130>
- 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>
- 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>
- 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>
- Wray, A.** (2008). *Formulaic language: Pushing the boundaries*. Oxford University Press.