Applied Psycholinguistics (2026), **47**, e22, 1–31
doi:10.1017/S0142716426100538

CAMBRIDGE UNIVERSITY PRESS

ORIGINAL ARTICLE

Learning binomials from context: Do non-natives develop offline and online sensitivity to novel binomials after exposure?

Suhad Sonbul[1], Gareth Carrol[2], Dina Abdel Salam El-Dakhs[3], Kathy Conklin[4] and Abdulaziz Altamimi[5]

[1]English Language Institute, Umm Al-Qura University, Saudi Arabia, [2]Department of Linguistics and Communication, University of Birmingham, UK, [3]Department of Linguistics and Translation, Prince Sultan University, Saudi Arabia, [4]School of English, University of Nottingham, UK and [5]Department of English Language and Literature, Imam Mohammad Ibn Saud Islamic University (IMSIU), Saudi Arabia
Corresponding author: Suhad Sonbul; Email: sssonbul@uqu.edu.sa

(Received 31 January 2025; revised 24 October 2025; accepted 11 February 2026)

Abstract

Research on how non-natives process and learn binomials (*black and white*) is limited. The present study addresses this gap using online (eye-tracking) and offline (familiarity rating) tasks. Sixty non-native speakers of English (L1 = Arabic) read six stories seeded with 21 novel binomials in three conditions: one exposure, six exposures, and no exposure (i.e., only in post-test) in a counter-balanced design. Each item was also presented in the reversed order (*white and black*). The non-natives read the stories as their eye movements were monitored and answered comprehension questions. In addition to the novel binomials, 12 existing binomials (congruent with Arabic) were included in the passages as a baseline for comparison. After completing the reading task, the participants completed an offline rating task as a measure of declarative knowledge of the binomial configuration (i.e., word order). All items were rated twice, once in the forward direction and once in the reversed direction. Online results showed that non-natives were not sensitive to the configuration of existing binomials, and there was limited evidence of any sensitivity to novel binomials. Offline, non-natives showed sensitivity to the configuration restrictions of existing binomials but not novel ones.

Keywords: Binomials; offline; online; L2 learning; non-native speakers

Introduction

English native speakers are more likely to say *salt and pepper* rather than *pepper and salt*, even though both phrases are grammatically acceptable and convey the same meaning. Lexical patterns, such as *salt and pepper,* are referred to as formulaic language and are ubiquitous in language, accounting for up to 50% of language use

© 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.

https://doi.org/10.1017/S0142716426100538 Published online by Cambridge University Press

(e.g., Erman & Warren, 2000). Previous research has demonstrated, using different methodologies, that English native speakers process formulaic language faster than novel language (e.g., Arnon & Snider, 2010; Carrol & Conklin, 2020; Siyanova-Chanturia et al., 2017; Tremblay et al., 2011). This processing advantage has been attributed to the relative frequency of formulaic sequences (compared to novel combinations), leading to their entrenchment in memory. The processing advantage observed for formulaic language in native speakers is not always evident in non-native speakers (e.g., Carrol & Conklin, 2014, 2017; Siyanova-Chanturia et al., 2011), and this has been attributed to their insufficient L2 (second language) exposure. Put another way, if repeated occurrences are needed to entrench a sequence in memory, non-natives may not have had sufficient exposure for this to occur.

Frequency of exposure has received increasing attention in vocabulary research. Previous research has demonstrated that frequency of exposure plays an important role in the development of L2 vocabulary knowledge (e.g., Godfroid et al., 2018; Laufer & Rozovski-Roitblat, 2011; Pellicer-Sánchez & Schmitt, 2010). However, the majority of studies have focused on single words, whereas comparatively fewer studies have examined formulaic sequences. The existing studies suggest that more exposure enhances the learning gains of L2 formulaic language, although most research has been carried out on collocations (Durrant & Schmitt, 2010; Majuddin et al., 2022; Webb et al., 2013). The aim of the present study is to investigate the role of exposure in the learning and processing of a less investigated type of formulaic language: binomials.

Binomials are defined as "phrases formed by two content words from the same lexical class connected by a conjunction, where one word order is more frequent than the other (*bride and groom*/*groom and bride*)" (Siyanova-Chanturia et al., 2011, p. 2). They differ from other types of formulaic sequences in several ways. For one, while binomials exhibit a preferred word order, altering the word order does not typically lead to changes in meaning or syntactic irregularity. This unique property, where word order can be reversed without creating ungrammatical formulations or introducing confounds (e.g., individual word frequency differences, syntactic and semantic oddities), makes binomials an ideal choice for examining the effect of phrasal frequency. Additionally, binomials allow the investigation of two key properties: co-occurrence restrictions (the combination of specific words) and configuration restrictions (the order of specific words).

Previous research on binomials has demonstrated a processing advantage for the conventionalized binomial form (e.g., *time and money*) relative to its reversed form (e.g., *money and time*) in online eye-tracking experiments (Siyanova-Chanturia et al., 2011). Sensitivity to the configuration restrictions of binomials has also been observed for novel binomials. Native speakers developed a processing advantage for novel binomials (e.g., *wires and pipes*) over their reversed forms (*pipes and wires*) after as few as four to five exposures (Conklin & Carrol, 2021) but this was not the case for non-native speakers (Sonbul et al., 2023). Research on binomial acquisition employing offline measures, on the other hand, has shown that non-native speakers can acquire knowledge of the preferred order of binomials after only two to six exposures (Alotaibi et al., 2022; Altamimi & Conklin, 2024b).

https://doi.org/10.1017/S0142716426100538 Published online by Cambridge University Press

As can be seen above, studies on binomials have relied on either online measures (eye-tracking) to capture real-time processing (Siyanova-Chanturia et al., 2011; Sonbul et al., 2023) or offline measures to assess declarative knowledge (Alotaibi et al., 2022) but rarely incorporated both within a single design. Crucially, research on formulaic language has demonstrated a dissociation between online and offline measures, suggesting that these measures tap into different knowledge systems (e.g., Sonbul & El-Dakhs, 2020; Sonbul & Schmitt, 2013; Sonbul, 2015). Another limitation of previous research on binomials is that the cross-language congruency effect has not been controlled in most previous studies (e.g., Sonbul et al., 2023).

The present study aims to contribute to research on the processing and learning of binomials by addressing these limitations and extending the work of Conklin and Carrol (2021) and Sonbul et al. (2023). More specifically, it examines both the learning and processing of binomials by incorporating online (eye-movements) and offline (familiarity rating) measures to determine whether sensitivity to the configuration of L2 binomials emerges incidentally during online reading as well as in offline post-test.

Literature review

Offline and online vocabulary measures

Most studies on the incidental acquisition of L2 vocabulary have focused on the acquisition of single words from reading exposure (e.g., Pellicer-Sánchez & Schmitt, 2010; Pitts et al., 1989; Waring & Takaki, 2003; Webb, 2007). These studies used offline post-tests to measure learning outcomes that are reported as gains in declarative knowledge as a result of increased exposure. Although offline measures have been commonly used in previous research, they have some key limitations. One concern is that they do not capture the "partial and incremental growth in knowledge occurring as a function of an individual's experience with each word" (Joseph et al., 2014, p. 245). To address this, some recent studies have started to explore vocabulary acquisition through online measures, such as eye-tracking. These methods allow for the examination of online processing as it unfolds in real time, offering insights into the incremental development of vocabulary knowledge with each repeated encounter during reading. Previous eye-tracking research suggests a link between eye-tracking patterns and vocabulary gains (e.g., Godfroid et al., 2013; Godfroid et al., 2018; Mohamed, 2018; Pellicer-Sánchez, 2016). These studies combined offline measures (paper-and-pencil post-tests) and online measures (eye-tracking) to investigate the incidental acquisition of L2 vocabulary. Godfroid et al. (2013) found that pseudowords (i.e., words that would be unknown to readers) had longer fixation times than known words for non-native learners. Further, longer fixations to the pseudowords were predictive of successful word recognition in a post-test. Later research by Godfroid et al. (2018) suggested that both the number of exposures and total reading time contributed to vocabulary learning and that reading times decreased as a result of increased exposure. Similarly, Pellicer-Sánchez (2016) showed that there was a positive relationship between total reading time and successful recall of pseudowords. In addition, the

https://doi.org/10.1017/S0142716426100538 Published online by Cambridge University Press

pseudowords elicited significantly faster reading times after three encounters, and by the eighth encounter, they were read in a similar manner to real words.

Other studies on formulaic language have revealed inconsistent findings in offline and online measures (e.g., Sonbul & El-Dakhs, 2020; Sonbul & Schmitt, 2013). Sonbul and Schmitt (2013) used both offline (i.e., form recall and recognition tasks) and online measures (i.e., a priming paradigm in a lexical decision task) to examine the learning of collocations. The offline measures indicated that all learning conditions enhanced knowledge of collocations for native and non-native speakers. In contrast, the online measure did not demonstrate a processing advantage for the target collocations. Sonbul and El-Dakhs (2020) employed both offline (i.e., a multiple-choice form recognition task) and online (i.e., a timed acceptability judgment task) measures to examine the recognition of two types of collocations: congruent (i.e., matching in form and meaning in English and Arabic) and incongruent collocations (e.g., existing only in English). Results from the offline task showed that Arabic-L1 learners of English-L2 recognized congruent collocations more accurately than incongruent ones. However, the online task revealed that response times were similar for both collocation types, with a modulating effect of proficiency whereby the difference was clearer for lower proficiency non-natives than higher proficiency ones. In a later eye-tracking experiment, El-Dakhs et al. (2024) found that non-natives process congruent and incongruent collocations similarly. In addition, a dissociation between online and offline measures was observed in Sonbul (2015), where sensitivity to collocational frequency appeared in offline measures (i.e., a rating task of how typical the collocations were in English). Online in an eye-tracking experiment, however, the effect was only evident in an early reading measure (first pass reading time) but not later in processing (i.e., total reading time and fixation count) for both native and non-native speakers.

One explanation for the discrepancy between offline and online measures is that they likely tap into different constructs. Offline measures assess declarative knowledge, whereas online measures reflect implicit processing. The findings from these studies suggest that even after declarative knowledge is acquired, it may not be accessible quickly enough to be detected during implicit processing tasks. It may be that more exposures to a lexical pattern are necessary before declarative knowledge is accessed automatically and can be detected by online measures.

The role of frequency in the learning and processing of vocabulary

Of all the factors that influence the learning and processing of single words, frequency of exposure has received the most attention in vocabulary research (Pellicer-Sánchez & Boers, 2019). Both offline and online measures have provided evidence for the role of frequency in learning. For example, previous studies using form recall and recognition post-tests have shown that frequency of exposure plays an important role in developing L2 vocabulary knowledge (e.g., Godfroid et al., 2018; Laufer & Rozovski-Roitblat, 2011; Pellicer-Sánchez & Schmitt, 2010), such that repeated contextual encounters with unknown L2 words contribute to learning gains. In addition, online measures have revealed that word frequency is one of the strongest predictors of response times, with more frequent words eliciting shorter response times than infrequent ones (Brysbaert et al., 2018). Important for the

https://doi.org/10.1017/S0142716426100538 Published online by Cambridge University Press

current study, eye-tracking research demonstrates that as learners encounter unknown words through repeated exposures during reading, these words elicit fewer and shorter fixations (e.g., Elgort et al., 2018; Godfroid et al., 2018; Pellicer-Sánchez, 2016). Thus, across repeated exposures (i.e., increasing frequency), reading times to new words become more like those of already known words, which is taken as an indication that the words are being learned.

In addition to its effect on single words, the role of frequency of exposure has also been shown to play an important role in the learning and processing of formulaic sequences. When measuring frequency effects, studies usually manipulate text-based frequency (i.e., the number of occurrences participants encounter) while controlling for corpus-based frequency. Research using offline measures indicates that number of encounters in a text contributes significantly to the acquisition of formulaic sequences from reading. Durrant and Schmitt (2010), for example, found that two exposures to a collocation led to better form recall by non-native learners than a single exposure. Webb et al. (2013) presented non-native learners with multiple instances (1-15 repetitions) of the same collocation in a text and found that the number of encounters positively influenced the incidental learning of collocations. More recently, Majuddin et al. (2022) exposed non-native learners to audiovisual input that was seeded with different types of formulaic sequences. Results showed that formulaic sequences that appeared twice yielded better recall than those that appeared only once. However, other studies have shown that increasing the number of exposures does not always lead to better learning gains. For example, Pellicer-Sánchez (2017) compared the acquisition of adjective-pseudoword collocations during reading under two exposure conditions: four versus eight times. Post-test results indicated incidental learning gains but revealed no significant effect based on the number of exposures.

Akin to what has been found for single words, we would expect that with repeated exposure to unknown formulaic language, processing will become more like that elicited by known formulaic sequences. Pellicer-Sánchez et al. (2022) investigated how frequency of exposure influences the processing of novel collocations. English native and non-native speakers read a story in which novel (adjective-pseudoword) collocations were presented either four or eight times, and known collocations were presented eight times. Results demonstrated an effect of frequency of exposure on collocation processing, with a decrease in both the number and duration of fixations across repeated exposures. Processing for novel collocations was comparable to that of known collocations after eight exposures for both native and non-native speakers. In another eye-tracking study, Altamimi and Conklin (2024a) replicated Pellicer-Sánchez et al.'s (2022) findings, demonstrating faster processing of novel collocations with repeated exposures among both native and non-native speakers.

The learning and processing of binomials by natives and non-natives

Most research on formulaic language has focused on either collocations or idioms, with comparatively fewer studies investigating the processing and acquisition of binomials. Binomials and collocations are similar in that learning them involves acquiring knowledge of the co-occurrences of words (e.g., *strong wind*, *salt and*

https://doi.org/10.1017/S0142716426100538 Published online by Cambridge University Press

pepper). However, they are different in that binomials require acquiring the additional knowledge of their specific configuration: developing a sensitivity to the word order preference (i.e., the forward form *salt and pepper* is much more frequent and natural sounding to a native speaker than the reversed form *pepper and salt*). In contrast, syntax dictates word order in collocations.

Phonological and semantic constraints have been proposed to account for why one word order is preferred in binomials (Benor & Levy 2006; Mollin, 2012). However, such accounts do not provide a complete explanation, as many counterexamples exist for each constraint (e.g., the male-before-female constraint as in *men and women* is violated in *bride and groom*) (Conklin & Carrol, 2021). Further, frequency was found to be a stronger predictor of word order preference than abstract constraints (Morgan & Levy, 2016). Thus, Conklin and Carrol (2021) suggest that the preferred word order of binomials is likely shaped by repeated usage, which leads to entrenchment of a particular order in memory, ultimately resulting in its conventionalization over time.

Previous research on binomial processing has capitalized on the syntactically permittable flexibility of word order for these sequences, demonstrating a processing advantage for the frequent conventionalized binomial form (*time and money*) relative to its reversed form (*money and time*), which differ only in phrasal frequency (e.g., Siyanova-Chanturia et al., 2011; 2017). Moreover, Siyanova-Chanturia et al. (2011) showed evidence that for non-native speakers of English, this sensitivity is modulated by L2 proficiency, with clearer effects as proficiency increases. The processing advantage for binomials over their reversed forms has been attributed to the entrenchment of the conventional pattern in memory.

More recently, studies have started to explore the processing of novel binomials. Conklin and Carrol (2021) used eye-tracking to investigate the processing of novel binomials among native speakers. Participants read short stories in which existing binomials were presented once in their conventional form (e.g., *time and money*), and novel binomials were presented one to five times in an experimentally defined forward form (e.g., *wires and pipes*). Subsequently, items were presented in their reversed forms (e.g., *money and time*, *pipes and wires*). Results demonstrated that, similar to existing binomials (which were always read more quickly than their reversed forms), novel binomials developed a processing advantage over their reversed forms after only four to five exposures. This suggests that sensitivity to the sequential order of binomials emerges rapidly during reading. Sonbul et al. (2023) replicated Conklin and Carrol's (2021) study with non-native learners (Arabic L1). They found no processing advantage for existing binomials over their reversed forms. In contrast to the original study, after either two or four repetitions of a novel binomial, the reversed forms were processed faster than their forward forms. This suggests that non-native speakers treated the reversed forms as another encounter of the same binomials, ignoring the configuration violation. That is, non-native speakers were able to register the co-occurrence restriction for novel binomials in memory but did not develop sensitivity to their sequential order (configuration) as quickly as native speakers.

Altamimi and Conklin (2024b) investigated the effects of congruency and frequency of exposure on the learning of binomials on self-paced reading and a timed word order recognition task. Arabic second language learners of English and native

https://doi.org/10.1017/S0142716426100538 Published online by Cambridge University Press

English speakers read a text in which three types of binomials were presented either two or five times: English-only, Arabic-only, and congruent (occurring in both English and Arabic). Results from the word order task showed that participants acquired recognition knowledge of the preferred order of binomials after only two exposures, with no effect of frequency of exposure. Analysis of response times showed that native English speakers recognized congruent and English-only items faster than Arabic-only items, while Arabic L2 learners recognized congruent items faster than both English-only and Arabic-only items. In terms of self-paced reading, there was a processing advantage across repeated exposures for both groups and under all item types.

Alotaibi et al. (2022) investigated the effects of the number of exposure (2, 4, 5, and 6 occurrences) and different input modes (reading-only, listening-only, and reading-while-listening) on the learning of binomials by non-native speakers. Post-test results (familiarity ratings) indicated that both reading-only and reading-while-listening were more effective for enhancing declarative knowledge of binomials than listening-only. There was also an effect of frequency: novel binomials encountered six times yielded familiarity ratings similar to existing binomials.

In sum, previous studies have shown that English native speakers demonstrate a clear sensitivity to the specific configuration of binomials. However, this is not always evident among non-native speakers. Further, the role of frequency of exposure on the learning and processing of binomials is still not fully understood.

The present study

Previous research suggests that frequency of exposure facilitates the learning and processing of formulaic sequences. The limited research on binomials has revealed discrepancies between native and non-native speakers' performance: for novel binomials, native speakers exhibit a sensitivity to both the co-occurrence of words and their configuration after repeated encounters—four to five exposures (Conklin & Carrol, 2021)—while non-native speakers seem to register the co-occurrence of words but not their configuration (Sonbul et al., 2023). The limitations in the Sonbul et al.'s (2023) study might account for such discrepancies. Sonbul et al. (2023) did not fully account for a cross-language congruency effect, which has been shown to influence the processing of formulaic sequences, particularly collocations and idioms (for an overview, see Conklin & Carrol, 2019). As indicated above, in addition to co-occurrence restrictions, which is a common feature of different types of formulaic sequences, binomials have a distinctive quality, that is, configuration restrictions. Hence, in binomials, congruency may have a different effect to other types of formulaic sequences. Additionally, Sonbul et al. (2023) included a limited range of repetitions: two versus four. As speculated by Sonbul et al. knowledge of configuration restrictions may be developed later than knowledge of co-occurrence restrictions, and thus, non-native speakers might require more exposures for sensitivity to configuration restrictions to emerge.

The aim of the present study is to address the limitations in Sonbul et al. (2023) and to extend its scope by accounting for the congruency effect, increasing the number of exposures, and including an offline post-exposure measure to assess declarative knowledge. This study seeks to answer the following research questions:

https://doi.org/10.1017/S0142716426100538 Published online by Cambridge University Press

RQ1: Do non-native speakers demonstrate online sensitivity to the configuration (i.e., word order) of existing congruent binomials and novel binomials encountered in context?

RQ2: How does the processing of novel binomials alter as a function of frequency of exposure (1 to 6 exposures)?

RQ3: Do non-native speakers exhibit offline sensitivity to the configuration of L1–L2 congruent existing binomials and novel binomials encountered in context?

Methods

Participants

A total of 61 participants (L1-Arabic L2-English) took part in the study. At the time of study, they worked or studied at a university in Saudi Arabia. One participant was excluded due to track failure during the experiment. Thus, we considered the data of 60 participants (Female = 39) who ranged in age between 18 and 42 (*Mean* = 23.58, *SD* = 5.39). Their self-rated L2 proficiency scores (out of five) in the four skills were as follows: speaking (*Mean* = 4.42, *SD* = 0.70), reading (*Mean* = 4.57, *SD* = 0.56), writing (*Mean* = 4.17, *SD* = 0.78), and listening (*Mean* = 4.18, *SD* = 0.70).

As an approximate objective measure of their proficiency in English, the participants completed the V_YesNo online vocabulary test (Meara & Miralpeix, 2017) with a maximum score of 10,000. Their average score was 6288,87 (*SD* = 1261.53, *Minimum* = 4092, *Maximum* = 9564). It should be noted that the vocabulary score was employed as a screening procedure in the present study whereby only participants who scored 4,000 or above completed the eye-tracking experiment and the offline task (see Procedures below). This was important to ensure that the materials (target/control items and passages, see below) were suitable for the participants and that they were not likely to experience any difficulty understanding them.

Since the study employed a three-version counter-balanced design (see Materials below), each participant was assigned one of the three versions: Version 1 (*N* = 20), Version 2 (*N* = 20), and Version 3 (*N* = 20). The analysis of a one-way ANOVA showed that participants reading the three versions did not differ in their vocabulary score: Version 1 (*Mean* = 7148.62, *SD* = 1560.59), Version 2 (*Mean* = 7018, *SD* = 1676.29), Version 3 (*Mean* = 6980.77, *SD* = 1514.12); (*df* = 2, *F* = 0.09, *p* =.91).

Items

The study set out to examine non-native learners' processing of existing English binomials that were congruent with Arabic and their processing/learning of novel English binomials. Hence, two sets of items were developed: existing English-Arabic binomials (same configuration) and novel binomials which do not exist in either language. All component words in the target and control binomials (existing and novel) belonged to the most frequent 4,000 word families in English (Nation, 2012).

https://doi.org/10.1017/S0142716426100538 Published online by Cambridge University Press

This helped to ensure that our participants, who all scored above 4,000 in the vocabulary test (see above), would be familiar with binomials' constituent words.

The existing binomials were developed based on extensive research in the literature (e.g., Altamimi, 2021; Benor & Levy, 2006; Conklin & Carrol, 2021; Gorgis & Al-Tamimi, 2005; Morgan & Levy, 2016). Selection of the items involved three steps. First, the frequency of the forward and backward forms of each potential binomial was checked in the NOW (News on the Web) corpus (Davies, 2016) with over 20 billion words. The NOW corpus is best suited to the purposes of the present study due to its considerable size (cf. the Corpus of Contemporary American English or COCA with just over a billion words). Bigger corpora allow for more accurate frequency counts for less frequent items, with binomials being less frequent in language use relative to other formulaic sequences such as collocations. Only items where one direction of the binomial was more frequent than the other with a minimum of 3:1 ratio (minimum forward raw frequency = 4,000) were included in the initial pool. This ensured that the binomial existed in English with a preferred direction. Second, two of the co-authors (proficient Arabic-English speakers) independently rated each of the 44 potential items for (1) whether the binomial is used in Arabic, and (2) whether the word order is the same in both English and Arabic. This resulted in a reduced list of 34 potential binomials, including only items that existed in Arabic and judged to be used in the same preferred direction as English. The next step was to check the Arabic frequency of the potential items in the 6.5-billion-word Arabic Web Corpus or "arTenTen" (Arts et al., 2014, available on Sketch Engine). Only items ($N = 21$) where the forward frequency was higher than the reversed frequency with a minimum ratio of 3:1 were included in the final pool of items (minimum forward raw frequency = 600).[1,2] Thus, the final list of existing binomials included 21 items with an average ratio of 31.04 to 1 in Arabic ($SD = 52.28$) and average ratio of 93.40 to 1 in English ($SD = 211.67$). The items had a word order preference that was significantly preferred and congruent in both languages: English (*Mean* forward = 67892.81, $SD = 95767.75$; *Mean* backward = 3850.81, $SD = 8627.84$; $t = 3.05$, $p = 0.006$) and Arabic (*Mean* forward = 15910.76, SD = 19522.45; *Mean* backward = 2214.48, $SD = 4005.86$; $t = 3.15$, $p = 0.005$). These items were divided into two categories: target items included in the reading texts ($N = 12$) and control items included in the offline rating task only ($N = 9$).

For novel binomials, similar procedures were followed. Some of the items were borrowed from Conklin and Carrol (2021), and additional items were constructed following the same procedures. "Noun and noun" word pairs were selected from the same semantic field with no clear order preference (e.g., *castle and tower*). The component words were selected in a way that none of them formed part of an existing frequent binomial. We ended up with an initial list of 44 items. First, we checked the potential novel binomials for their English frequency in NOW. Only items with a low frequency in either direction (fewer than 2,000 raw occurrences) and no preferred word order in English (forward-to-backward ratio of less than 2 to 1) were selected. This resulted in a reduced list of 40 items. Finally, we checked the Arabic frequency of the novel binomials in both directions in the "arTenTen" Corpus and only included items where the frequency was very low (less than 500 occurrences with a forward-backward and backward-forward ratio of less than 2:1).[3] For the final list of novel items, one direction was specified as the experimentally

https://doi.org/10.1017/S0142716426100538 Published online by Cambridge University Press

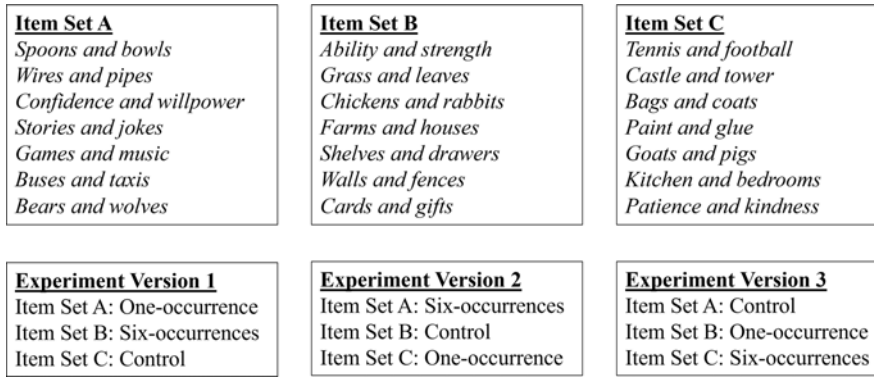


Figure 1. Counterbalancing under the three versions of the experiment.

defined forward direction and the other one was operationalized as the backward form. The novel items did not differ in the forward or backward direction either in English (*Mean* forward = 321.33, *SD* = 341.87; *Mean* backward = 313.14, *SD* = 269.51; $t = 0.09$, $p = .93$) or Arabic (*Mean* forward = 82.52, *SD* = 181.57; *Mean* backward = 158.86, *SD* = 450.06; $t = -0.72$, $p = 0.48$). Appendix S1 in the online supplementary material presents the existing and novel binomials and their properties. The items were divided into three sets (A, B, C) counter-balanced across the three conditions: One occurrence in the reading texts (1-rep henceforth), six occurrences in the reading texts (6-rep henceforth), and no occurrences in the reading texts (control henceforth) (see Figure 1). The no-occurrence items will be included in an offline rating task following the eye-tracking experiment as control items (see further below).

Reading texts

Six stories were developed by the second author (a native speaker of English) to reflect the counter-balanced versions presented in Figure 1. Items were distributed proportionally across the six stories as follows:

1. For the existing binomials, the forward and reversed forms of each binomial appeared in consecutive stories (e.g., forward in Story 1 and reversed in Story 2; forward in Story 2 and reversed in Story 3, etc.).
2. For the novel binomials which appeared only once (1-rep condition), the forward and reversed forms appeared in consecutive stories in the same way as for existing binomials.
3. For the novel binomials which appeared six times (6-rep condition), the six forward forms appeared in four consecutive stories in a 2-1-2-1 design (e.g., two occurrences in Story 1—1 occurrence in Story 2—2 occurrences in Story 3—1 occurrence in Story 4). This was finally followed by the reversed form in the next story (e.g., Story 5). All 6-rep items therefore began in either Story 1 or 2 and ended in Story 5 or 6.

https://doi.org/10.1017/S0142716426100538 Published online by Cambridge University Press

4. For the novel binomials in the control condition, these did not appear at all in the reading passages.

Following the recommendations of Arai and Takizawa (2024), we checked the lexical coverage of the passages using the VocabProfiler Tool on LexTutor.com. Results showed that 99% of the words in the texts belonged to the most frequent 4,000 word families in English, which our participants were familiar with (see *Participants* above).

For each story, three yes/no comprehension questions were created to ensure that participants were reading for comprehension (Total = 18 questions). The full passages, comprehension questions, data files, and codes are available on OSF through the following link: https://osf.io/x4whb.

Familiarity rating task

As a measure of offline processing (declarative knowledge) of the existing and novel binomials, participants completed a familiarity rating task. In this task, the items (Existing = 21, Novel = 21, Total = 42) were randomly presented in their forward and reversed forms to participants in two counter-balanced lists with a balanced number of items in both directions under each list. No item was presented in its forward and reversed forms in the same list. The items belonged to one of the following categories:

1- Existing binomials which were presented during the eye-tracking experiment ($N = 12$).
2- Existing binomials which were not presented during the eye-tracking experiment ($N = 9$).
3- Novel binomials presented during the eye-tracking experiment once (1-rep type) ($N = 7$)
4- Novel binomials presented during the eye-tracking experiment six times (6-rep type) ($N = 7$)
5- Novel binomials not presented during the eye-tracking experiment (control type) ($N = 7$)

The instructions were as follows:

> In this task, you will see a set of English phrases in isolation, for example "black and white." You need to decide on a scale of 1 to 7 how familiar each is. On this scale, 1 = very unfamiliar and 7 = very familiar. For example, for the phrase "gold and silver," I would circle "7" because I have heard this phrase a lot of times. However, "silver and gold," although using the same words would be much less familiar so would receive lower ratings, probably around 1 to 3.

Procedure

The participants were invited to the eye-tracking lab, and informed consent was obtained. The study started with the V_YesNo online vocabulary test as a rough

https://doi.org/10.1017/S0142716426100538 Published online by Cambridge University Press

measure of proficiency (15 minutes). Participants who crossed the threshold score of 4,000 in this test were invited to complete the eye-tracking experiment. Next, the eye-tracking experiment started with a practice story, and then, the six main stories appeared sequentially, each followed by three comprehension questions. Finally, the participants completed the rating task and answered some questions about their language background.

The eye-tracking experiment was run on SR Research Eye-link 1000+ eye-tracker in monocular mode at 500Hz with a desk-mounted chin rest to minimize movement. A 9-point grid calibration was conducted before each story and whenever it was deemed necessary during the experiment, and a fixation point preceded each screen for drift checking. The stories were presented in 18-point Courier New font, double-spaced. The task was to read the stories as naturally as possible for comprehension and to answer the questions that followed each story. Target binomials did not appear in the beginning or end of lines or at line breaks. Results of the comprehension questions showed that participants did not have any issues understanding the stories (*Mean* = 88.43%, *SD* = 10.64, *Minimum* = 61.11%, *Maximum* = 100%).

Analysis

The analysis was performed using linear mixed-effects models in R 4.4.2 (R Core Team, 2024). We carried out four analyses to answer the three research questions: Analysis 1 (RQ1) and Analysis 2 (RQ2) analyzed online eye-tracking data while Analysis 3 and Analysis 4 examined offline familiarity ratings (RQ3). All analyses were run using the *lmer4* and *lmerTest* functions. For the eye movement data (Analysis 1 and Analysis 2), we examined first-pass reading time (FPRT) and total reading time (TRT) as measures of early and late processing, respectively. These two measures are often employed in research on formulaic language (e.g., Conklin & Carrol, 2021) as they can be viewed as being complementary. While FPRT exhibits lexical-level processing of the target items, TRT incorporates textual integration considering the wider context (including regressions to the target items). Two regions of interest were examined: Word 3, and whole phrase.[4] Since the word "and" was skipped more than 60% of the time, it was not considered as an individual region of analysis.

We started by cleaning the eye-tracking data using the four-step procedure; fixations shorter than 100 ms and longer than 800 ms were removed (4.2% of all fixations). We excluded trials that were discontinued or where track failure was experienced. Also, a phrase that was completely skipped was also excluded from the analysis. In such cases, all subsequent occurrences of the item, including the reversed form, were also removed. This resulted in loss of 14.6% of data points in both Analysis 1 and Analysis 2.[5] Both reading time measures (TRT and FPRT) were log-transformed prior to the analysis to reduce skewness in the data.

Analysis 1 examined the eye-movement data for the forward (first occurrence) and reversed forms of both the existing and novel binomials to answer RQ 1 (Do non-native speakers demonstrate online sensitivity to the configuration (i.e., word order) of existing congruent binomials and novel binomials encountered in context?). The outcome variables were FPRT and TRT for the two main regions of

https://doi.org/10.1017/S0142716426100538 Published online by Cambridge University Press

interest; hence, four models were fitted. The analysis used the maximal model justified by the research design. Fixed effects included Item Type (existing, 1-rep, and 6-rep) and Direction (forward versus reversed), and the interaction between them. We controlled for several variables to partial out their effect, and these included vocabulary size, Word 1 frequency, Word 3 frequency, Word 1 length, and Word 3 length. Frequency was based on the NOW Corpus and was expressed on the Zipf scale (van Heuven et al., 2014), and length was counted in characters. Then, we added several factors to examine their contribution to the model fit: forward association strength (Kiss et al., 1973), phrasal frequency, and the Item Type x Direction x Vocabulary score interaction. Only variables that significantly improved the model fit (with lower AIC values) were retained. Following von der Malsburg and Angele's (2017) suggestions for corrections in eye-tracking experiments that include more than one measure, we divided the .05 threshold by the number of measures (i.e., two: FPRT and TRT), and only treated a variable as significant if it reached the adjusted threshold of .025. By-participant slopes for Item Type and Direction were initially added. However, including them both resulted in convergence issues, so only Item Type was included as a slope. When the Item Type x Direction was significant, pairwise comparisons were conducted using the *emmeans* function in the *emmeans* package. A significant difference between the forward and reversed forms under one of the three item types (existing, 1-rep, and 6-rep) was taken as evidence for sensitivity to binomial configuration.

Analysis 2 focused on the 6-rep novel items only to answer RQ2 (How does the processing of novel binomials alter as a function of frequency of exposure (1 to 6 exposures)?). The main fixed factor in this analysis was occurrence (first, second, third, fourth, fifth, sixth, and reversed). The covariates were the same as Analysis 1, with the addition of Item Set which was included to account for the effect of counterbalancing. Similar to Analysis 1, we examined the contribution of forward association, phrasal frequency, and the Vocabulary score x Occurrence interaction. Only factors that significantly improved the model fit were included in the final model.

Finally, Analysis 3 and Analysis 4 examined the rating data from the familiarity task for the existing and novel binomials, respectively, to answer RQ3 (Do non-native speakers exhibit offline sensitivity to the configuration of L1-L2 congruent existing binomials and novel binomials encountered in context?). For offline results, the significance threshold was set at 0.05. The fixed factors in Analysis 3 for existing binomials included Condition (target versus control), Direction (forward versus reversed), and the interaction between them. When a significant difference emerges between the forward and reversed forms under either condition, this can be taken as evidence of offline sensitivity to the configuration of existing binomials. An additional aim here was to compare targets seen during the eye-tracking experiment with those that were not (controls), to establish any differences in ratings as a result of exposure during reading.

For Analysis 4, the fixed factors included Condition (1-rep, 6-rep, and control), Direction (forward versus reversed), and the interaction between them. Significant differences between the forward and reversed forms under any of the conditions were taken as evidence of non-natives' developing offline sensitivity to novel binomials after exposure. Covariates were the same as those examined in the

https://doi.org/10.1017/S0142716426100538 Published online by Cambridge University Press

previous analyses, including the Condition x Direction x Vocabulary score interaction. By-participant slopes for Condition and Direction were included in both analyses, but this resulted in convergence issues, so only Condition was included in the final models.

For all models that were fitted, we checked for collinearity using the VIF (variance inflation factor) values and no issues emerged (all VIF values < 7).

Results

Analysis 1: RQ1 (Do non-native speakers demonstrate online sensitivity to the configuration (i.e., word order) of existing congruent binomials and novel binomials encountered in context?)

Table 1 presents the average FPRT and TRT (raw data) for existing and novel binomials in the forward (first occurrence) and reversed forms. The best-fit models for Analysis 1 (based on log-transformed data) are presented in Table 2. Results showed that the Item Type x Direction interaction was significant for both measures and for all interest areas. Pairwise comparisons (see Appendix S2, online supplementary material, and Figure 2) showed that the reversed form of novel binomials in the 6-rep condition was always read significantly faster than the forward form (between 53 and 192 ms difference). Additionally, for the 1-rep condition, the reversed form showed faster processing than the forward form during late processing (TRT) only for the whole phrase (100 ms difference). As for the existing binomials, the difference between the forward and reversed forms was never significant.[6]

Analysis 2: RQ2 (How does the processing of novel binomials alter as a function of frequency of exposure (1 to 6 exposures)?)

Analysis 2 examined the FPRT and TRT of the 6-rep novel items to understand the effect of frequency of exposure (1–6) on online processing. Mean reading times are presented in Table 3, and the models are presented in Table 4. Pairwise comparisons are included in Appendix S3 in the online supplementary material (see also Figure 3). Results showed that occurrence had a significant effect on reading times (early and late measures) for Word 3 and the whole phrase. Pairwise comparisons showed that the first occurrence of binomials was almost always read more slowly than all the other occurrences (including the reversed forms). This seems to suggest that reading a binomial multiple times led to a processing advantage. The fact that the reversed form was read faster than the initial one supports the finding in Analysis 1 whereby the reversed 6-rep forms were read faster than the initial forward exposure.

It is evident that certain comparisons of the occurrences were significant (but not all) with a positive value, suggesting that more exposure (even the reversed form) led to faster reading, though not always significantly so. The overall downward trend in Figure 3 supports this finding.

https://doi.org/10.1017/S0142716426100538 Published online by Cambridge University Press

https://doi.org/10.1017/S0142716426100538 Published online by Cambridge University Press

Table 1. Average reading times (and standard deviation) for existing and novel items in the forward and reversed directions

Measure	Region of interest	Existing binomial			Novel binomial (1-repetition)			Novel binomial (6-repetition)		
		Forward	Reversed	Reversed-Forward	Forward	Reversed	Reversed-Forward	Forward	Reversed	Reversed-Forward
First pass reading time	Whole phrase	697.31	678.17	−19.14	844.82	768.59	−76.23	841.7	698.74	−142.96
		(403.91)	(387.1)		(511.64)	(437.76)		(501.45)	(357.77)	
	Word 3	301.72	319.44	17.72	392.61	364.84	−27.77	382.68	329.35	−53.33
		(153.07)	(142.66)		(234.34)	(178.91)		(199.54)	(142.83)	
Total reading time	Whole phrase	885.02	878.14	−6.88	1054.84	954.74	−100.1	1058.26	866.04	−192.22
		(467.42)	(455.2)		(590.03)	(577.4)		(574.51)	(448.24)	
	Word 3	354.46	372.13	17.67	464.39	409.8	−54.59	469.66	374.21	−95.45
		(207.8)	(209.44)		(306.91)	(233.4)		(294.65)	(200.22)	

https://doi.org/10.1017/S0142716426100538 Published online by Cambridge University Press

Table 2. Model 1 for the difference between forward and reversed forms for existing and novel binomials

	First pass reading time[a]					Total reading time[b]				
	Estimate	SE	t value	*p*	Cohen's d	Estimate	SE	t value	*p*	Cohen's d
Whole phrase										
(Intercept)	6.38	4.90	129.51	<.001***	10.30	6.62	0.05	137.78	<.001***	14.01
Log vocabulary test score	−0.18	3.50	−5.08	<.001***	−0.29	−0.21	0.04	−5.39	<.001***	−0.46
W1 Length	0.03	1.60	1.80	.09	0.05	0.05	0.02	3.04	.005**	0.11
W3 Length	0.03	1.60	1.65	.11	0.04	0.04	0.02	2.26	0.03	0.08
W1 Zip frequency	−0.04	1.70	−2.40	.023*	−0.07	−0.04	0.02	−2.58	.015*	−0.10
W3 Zip frequency	0.00	1.80	0.04	.97	0.00	−0.01	0.02	−0.59	0.56	−0.02
Item Type 1-rep	0.03	5.10	0.63	.53	0.05	0.09	0.04	2.29	0.026	0.19
Item Type 6-rep	0.06	5.10	1.20	.24	0.10	0.10	0.04	2.77	.008**	0.22
Direction reverse	−0.01	2.90	−0.43	.67	−0.02	0.00	0.02	−0.06	0.95	0.00
Forward association	−0.04	1.70	−2.57	.018*	−0.07	–	–	–	–	–
Item Type 1-rep: Direction reverse	−0.05	4.80	−1.12	.26	−0.09	−0.11	0.03	−3.35	<.001***	−0.23
Item Type 6-rep: Direction reverse	−0.14	4.80	−3.00	.003**	−0.23	−0.19	0.03	−5.78	<.001***	−0.39
Random effects	Variance	SD				Variance	SD			
Subject	0.06	0.24				0.06	0.24			
Subject | Item Type 1-rep	0.01	0.10				0.01	0.10			
Subject | Item Type 6-rep	0.01	0.09				0.01	0.09			
Item	0.00	0.05				0.00	0.05			
Residual	0.30	0.55				0.30	0.55			

(*Continued*)

https://doi.org/10.1017/S0142716426100538 Published online by Cambridge University Press

Table 2. (*Continued*)

	First pass reading time[a]					Total reading time[b]				
	Estimate	SE	t value	*p*	Cohen's d	Estimate	SE	t value	*p*	Cohen's d
Word 3										
(Intercept)	5.59	0.03	189.43	<.001***	13.08	5.72	0.04	156.06	<.001***	11.93
Log vocabulary test score	−0.10	0.02	−4.83	<.001***	−0.24	−0.12	0.03	−4.60	<.001***	−0.26
W1 Length	0.02	0.01	1.31	0.20	0.04	0.03	0.02	1.77	0.09	0.06
W3 Length	0.03	0.01	2.23	0.03	0.07	0.04	0.02	2.28	0.03	0.08
W1 Zip frequency	−0.02	0.01	−1.63	0.11	−0.06	−0.02	0.02	−1.39	0.17	−0.05
W3 Zip frequency	−0.02	0.01	−1.45	0.16	−0.05	−0.03	0.02	−1.54	0.14	−0.06
Item Type 1-rep	0.18	0.03	5.33	<.001***	0.42	0.18	0.04	4.33	<.001***	0.37
Item Type 6-rep	0.17	0.03	5.12	<.001***	0.41	0.20	0.04	5.06	<.001***	0.42
Direction reverse	0.06	0.02	2.66	.008**	0.14	0.05	0.02	2.04	0.04	0.11
Item Type 1-rep: Direction reverse	−0.11	0.04	−2.94	.003**	−0.25	−0.14	0.04	−3.51	<.001***	−0.30
Item Type 6-rep: Direction reverse	−0.18	0.04	−4.94	<.001***	−0.43	−0.24	0.04	−5.98	<.001***	−0.50
Random effects	Variance	SD				Variance	SD			
Subject	0.01	0.12				0.02	0.16			
Subject | Item Type One	0.00	0.01				0.00	0.05			
Subject | Item Type Six	0.00	0.03				0.00	0.04			
Item	0.00	0.05				0.00	0.06			
Residual	0.16	0.41				0.20	0.44			

[a]Analysis models:Whole phrase: *(lg_IA_FIRST_RUN_DWELL_TIME ~ lg_VocabularyTestScore + W1Length + W3Length + W1ZipFrequency + W3ZipFrequency + ItemType*Direction + ForwardAssociation+ (1+ItemType|Subject) + (1|Item)).*
Word 3: *(lg_IA_FIRST_RUN_DWELL_TIME ~ lg_VocabularyTestScore + W1Length + W3Length + W1ZipFrequency + W3ZipFrequency + ItemType*Direction + (1+ItemType|Subject) + (1|Item)).*
[b]Analysis models:
Whole phrase: *(lg_IA_DWELL_TIME ~ lg_VocabularyTestScore + W1Length + W3Length + W1ZipFrequency + W3ZipFrequency + ItemType*Direction + (1+ItemType|ParticipantCode) + (1|Item)).*
Word 3: *(lg_IA_DWELL_TIME ~ lg_VocabularyTestScore + W1Length + W3Length + W1ZipFrequency + W3ZipFrequency + ItemType*Direction + (1+ItemType|Subject) + (1|Item)).*

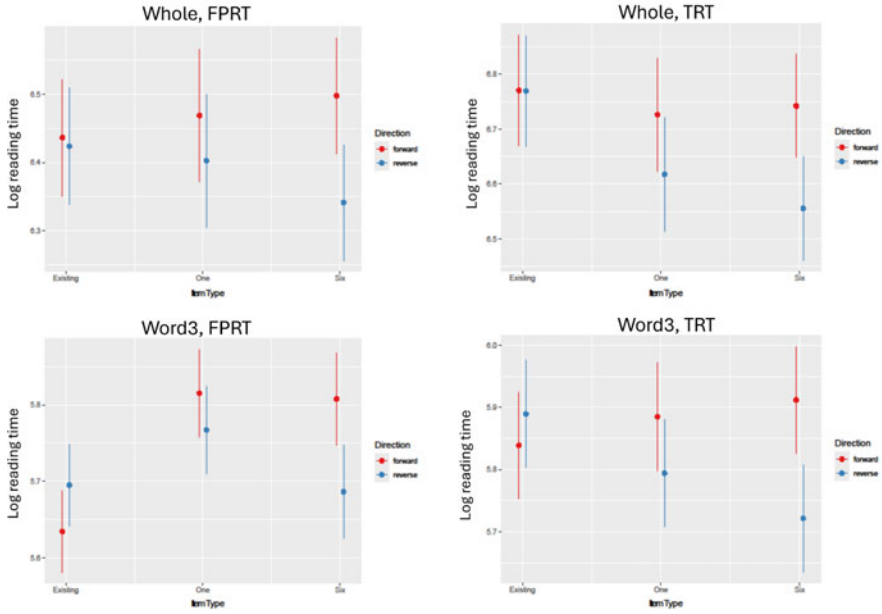


Figure 2. A graphical presentation of the Item Type x Direction interaction (Analysis 1).

Analysis 3 & Analysis 4: RQ3 (Do non-native speakers exhibit offline sensitivity to the configuration of L1-L2 congruent existing binomials and novel binomials encountered in context?)

Mean rating scores for existing and novel binomials are presented in Table 5, and the best-fit model for existing binomials (Analysis 3) is presented in Table 6. Results showed a significant interaction between Condition and Direction. Pairwise comparisons (Table 7) indicated a significant increase in rating for the forward forms of binomials over their reversed forms for existing binomials under both the target (presented during reading, 1.29 difference) and control (not presented during reading, 0.7 difference) conditions. This suggests that L2 participants had declarative knowledge of the preferred word order of English binomials that are congruent with Arabic. Results also showed that the forward forms were rated higher for items presented during reading (target) than those that were not (control), but this was not the case for the reversed forms. This may represent a recency effect due to exposure during the reading task.

Turning now to the novel binomials (Analysis 4), the best-fit model is presented in Table 8. Results showed a main effect of Condition but no effect of Direction and no interaction between the two factors. Thus, target novel binomials were rated higher in familiarity than control novel binomials. This result seems to suggest that regardless of direction, novel binomials that were presented during reading (once or six times) were rated higher than control items which were not included in the stories. Since Direction was not significant, it can be concluded that non-natives do not seem to have developed a sensitivity to novel binomials' configuration

https://doi.org/10.1017/S0142716426100538 Published online by Cambridge University Press

https://doi.org/10.1017/S0142716426100538 Published online by Cambridge University Press

Table 3. Mean reading times (standard deviation) for the various repetitions of the novel items that were repeated six times

		Occurrence 1 (Forward)	Occurrence 2	Occurrence 3	Occurrence 4	Occurrence 5	Occurrence 6	Reversed
First pass reading time	Whole phrase	841.7	752.31	756.93	729.52	658.26	708.03	698.74
		(501.45)	(456.77)	(420)	(430.7)	(367.59)	(395.83)	(357.77)
	Word 3	382.68	334.53	350.21	342.99	318.67	335.29	329.35
		(199.54)	(229.38)	(176.31)	(174.36)	(165.66)	(157.25)	(142.83)
Total reading time	Whole phrase	1058.26	917.19	934.28	935.34	848.82	886.74	866.04
		(574.51)	(525.97)	(477.01)	(486.13)	(442.81)	(466.33)	(448.24)
	Word 3	469.66	384.79	402.42	388.86	366.18	374.32	374.21
		(294.65)	(292.22)	(238.37)	(219.66)	(225.1)	(208.8)	(200.22)

https://doi.org/10.1017/S0142716426100538 Published online by Cambridge University Press

Table 4. Model 2 for the difference between the various repetitions of the 6-repetition novel items

	First pass reading time[a]					Total reading time[b]				
	Estimate	SE	t value	*p*	Cohen's d	Estimate	SE	t value	*p*	Cohen's d
Whole phrase										
(Intercept)	5.93	0.06	93.45	<.001***	12.15	6.79	0.07	93.51	<.001***	14.83
Item Set B	−0.08	0.08	−0.96	.34	−0.16	−0.10	0.10	−1.02	0.31	−0.21
Item Set C	−0.11	0.08	−1.30	.20	−0.22	−0.10	0.10	−1.08	0.28	−0.23
Log vocabulary test score	−0.17	0.03	−5.26	<.001***	−0.34	−0.20	0.04	−4.94	<.001***	−0.43
W1 Length	0.05	0.04	1.33	.21	0.10	0.03	0.03	0.82	0.42	0.06
W3 Length	−0.01	0.03	−0.21	.84	−0.01	0.01	0.03	0.39	0.71	0.03
W1 Zip frequency	−0.04	0.02	−1.71	.11	−0.08	−0.04	0.02	−1.90	0.08	−0.08
W3 Zip frequency	−0.03	0.03	−0.99	.34	−0.05	−0.04	0.02	−1.61	0.13	−0.08
Occurrence 2	−0.12	0.03	−3.68	<.001***	−0.24	−0.14	0.03	−5.62	<.001***	−0.31
Occurrence 3	−0.06	0.03	−1.84	.07	−0.12	−0.12	0.03	−4.92	<.001***	−0.27
Occurrence 4	−0.07	0.03	−2.11	.04	−0.13	−0.12	0.03	−4.80	<.001***	−0.27
Occurrence 5	−0.20	0.03	−6.47	<.001***	−0.41	−0.21	0.03	−8.44	<.001***	−0.47
Occurrence 6	−0.12	0.03	−3.79	<.001***	−0.24	−0.17	0.03	−6.70	<.001***	−0.37
Occurrence reversed	−0.15	0.03	−4.88	<.001***	−0.31	−0.19	0.03	−7.36	<.001***	−0.41
Random effects	Variance	SD				Variance	SD			
Subject	0.04162	0.204				0.07	0.27			
Item	0.0061	0.0781				0.00	0.07			
Residual	0.19047	0.4364				0.13	0.37			

(*Continued*)

https://doi.org/10.1017/S0142716426100538 Published online by Cambridge University Press

Table 4. (*Continued*)

	First pass reading time[a]					Total reading time[b]				
	Estimate	SE	t value	*p*	Cohen's d	Estimate	SE	t value	*p*	Cohen's d
Word 3										
(Intercept)	5.82	0.05	123.15	<.001***	13.52	5.97	0.05	109.90	<.001***	12.44
Item Set B	−0.07	0.06	−1.11	.27	−0.15	−0.11	0.07	−1.63	.11	−0.22
Item Set C	−0.10	0.06	−1.67	.10	−0.23	−0.18	0.07	−2.69	.0097**	−0.37
Log vocabulary test score	−0.11	0.02	−4.77	<.001***	−0.27	−0.13	0.03	−4.87	<.001***	−0.27
W1 Length	−0.05	0.03	−1.78	.10	−0.11	−0.08	0.03	−2.51	.03	−0.16
W3 Length	0.07	0.02	2.94	.01*	0.16	0.10	0.03	3.68	.003**	0.20
W1 Zip frequency	−0.03	0.02	−1.73	.11	−0.06	−0.01	0.02	−0.69	.50	−0.03
W3 Zip frequency	−0.05	0.02	−3.05	.009**	−0.13	−0.07	0.02	−3.53	.004**	−0.14
Occurrence 2	−0.15	0.03	−5.35	<.001***	−0.35	−0.21	0.03	−6.64	<.001***	−0.44
Occurrence 3	−0.08	0.03	−2.64	.008**	−0.17	−0.14	0.03	−4.33	<.001***	−0.29
Occurrence 4	−0.10	0.03	−3.54	<.001***	−0.23	−0.17	0.03	−5.27	<.001***	−0.35
Occurrence 5	−0.17	0.03	−6.12	<.001***	−0.41	−0.24	0.03	−7.51	<.001***	−0.50
Occurrence 6	−0.11	0.03	−4.03	<.001***	−0.27	−0.20	0.03	−6.22	<.001***	−0.41
Occurrence reversed	−0.12	0.03	−4.18	<.001***	−0.28	−0.19	0.03	−5.94	<.001***	−0.39
Phrase Zip frequency	–	–	–	–	–	−0.09	0.03	−2.58	.023*	−0.19
Random effects	Variance	SD				Variance	SD			
Subject	0.02	0.15				0.03	0.17			
Item	0.00	0.05				0.00	0.05			
Residual	0.16	0.40				0.20	0.45			

[a]Analysis models:Whole phrase: *(lg_IA_FIRST_RUN_DWELL_TIME ~ ItemSet + lg_VocabularyTestScore + W1Length + W3Length + W1ZipFrequency + W3ZipFrequency + Occurrence + (1|Subject) + (1|Item)).*

Word 3: *(lg_IA_FIRST_RUN_DWELL_TIME ~ ItemSet + lg_VocabularyTestScore + W1Length + W3Length + W1ZipFrequency+ W3ZipFrequency + Occurrence + (1|Subject) + (1|Item)).*

[b]Analysis models:

Whole phrase: *(lg_IA_DWELL_TIME ~ ItemSet + lg_VocabularyTestScore + W1Length + W3Length + W1ZipFrequency + W3ZipFrequency + Occurrence + (1|Subject) + (1|Item)).*

Word 3: *(lg_IA_DWELL_TIME ~ ItemSet + lg_VocabularyTestScore + W1Length + W3Length + W1ZipFrequency + W3ZipFrequency + Occurrence + PhraseZipFrequency + (1|Subject) + (1|Item)).*

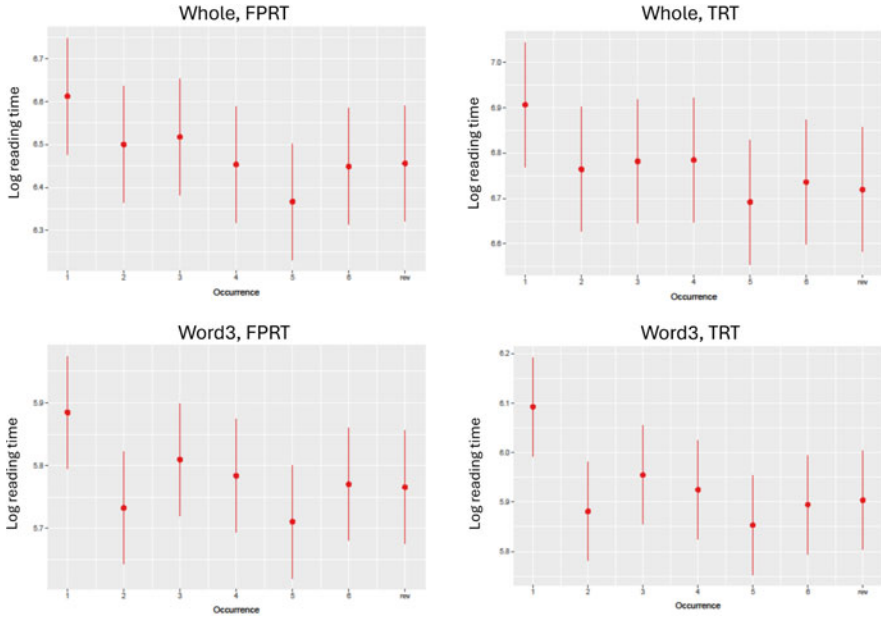


Figure 3. A graphical presentation of "Occurrence" on reading times (Analysis 2).

Table 5. Means (standard deviation) for the rating (scale: 1 = very unfamiliar to 7 = very familiar) of existing and novel binomials under the various conditions

Item Type	Condition[a]	Forward	Reversed	Difference
Existing binomials	Control	5.34	4.64	0.7
		(2.34)	(2.51)	
	Target	6.11	4.82	1.29
		(1.81)	(2.5)	
Novel binomials	Control	4.53	4.55	−0.02
		(2.33)	(2.36)	
	Target-1rep	5.07	5.21	−0.14
		(2.16)	(2.11)	
	Target-6 rep	5.8	5.75	0.05
		(1.9)	(1.84)	

[a]Control items were not included in the reading passages, but target items were included in the reading passages.

https://doi.org/10.1017/S0142716426100538 Published online by Cambridge University Press

Table 6. Model for the difference between forward and reversed forms for the existing binomials (Analysis 3) in the offline rating task

	Estimate	*SE*	*df*	*t value*	*p*
(Intercept)	5.19	0.31	74.34	16.76	<.001***
Log vocabulary test score	0.17	0.21	58.00	0.79	.43
W1 Length	0.01	0.05	14.00	0.28	.79
W3 Length	−0.02	0.05	14.00	−0.40	.69
W1 Zip frequency	−0.06	0.08	14.00	−0.67	.51
W3 Zip frequency	0.01	0.08	14.00	0.14	.89
Condition: Target	0.82	0.20	66.21	4.15	<.001***
Direction: reverse	−0.70	0.10	2379.00	−7.04	<.001***
Phrase Zip frequency	0.24	0.13	14.00	1.83	.09
Condition One: Direction reverse	−0.59	0.13	2379.00	−4.45	<.001***
Random effects	Variance	SD			
Subject	3.47	1.86			
Subject | Condition: Target	1.35	1.16			
Item	0.02	0.13			
Residual	2.70	1.64			

Analysis model: (RatingScore ~ lg_VocabularyTestScore + W1Length + W3Length + W1ZipFrequency + W3ZipFrequency + Condition*Direction + PhraseZipFrequency + (1+Condition|Subject) + (1|Item)).

Table 7. Pairwise comparison for the Condition x Direction interaction (Existing binomials – offline task)

		Estimate	*SE*	*df*	*t value*	*p*	*Cohen's d*
Comparison by direction							
Target Condition	forward-reversed	1.29	0.09	2379	14.92	<.001***	0.47
Control Condition	forward-reversed	0.704	0.1	2379	7.04	<.001***	0.26
Comparison by Condition							
Forward	Control-target	−0.82	0.20	66.20	−4.15	<.001***	−0.30
Reversed	Control-target	−0.24	0.20	66.20	−1.19	0.64	−0.09

restrictions after exposure. We also examined the 1-rep versus 6-rep difference and found that the former was rated lower than the latter (*estimate* = −0.63, $t = -4.95$, $p < .001$), indicating an overall increase in familiarity as a result of more exposures.

https://doi.org/10.1017/S0142716426100538 Published online by Cambridge University Press

Table 8. Model for the difference between forward and reversed forms for the novel binomials (Analysis 4) in the offline rating task

	Estimate	*Std. Error*	*df*	*t value*	*p*
(Intercept)	4.80	0.29	84.29	16.49	<.001***
Item Set B	−0.24	0.20	18.59	−1.23	.23
Item Set C	−0.36	0.20	18.39	−1.79	.09
Log vocabulary test score	−0.01	0.19	57.94	−0.08	.94
W1 Length	0.04	0.15	14.00	0.25	.81
W3 Length	0.06	0.15	14.00	0.42	.68
W1 Zip frequency	0.07	0.08	14.00	0.88	.39
W3 Zip frequency	0.07	0.09	14.00	0.74	.47
Condition TaregtOne	0.53	0.16	95.71	3.43	<.001***
Condition TargetSix	1.27	0.24	71.65	5.35	<.001***
Direction reverse	0.01	0.11	2319.00	0.14	.89
Condition One: Direction reverse	0.13	0.15	2319.00	0.88	.38
Condition Six: Direction reverse	−0.07	0.15	2319.00	−0.47	.64
Random effects	Variance	SD			
Subject	3.11163	1.764			
Subject | Condition: Target One	0.78784	0.8876			
Subject | Condition: Target Six	2.70648	1.6451			
Item	0.07943	0.2818			
Residual	2.32015	1.5232			

Analysis model: (RatingScore ~ ItemSet + lg_VocabularyTestScore + W1Length + W3Length + W1ZipFrequency + W3ZipFrequency + Condition*Direction + (1+Condition|Subject) + (1|Item)).

Discussion

The primary goal of this study was to determine whether non-natives demonstrate an offline or online sensitivity to configuration constraints in novel binomials after six repetitions. The study combined an online measure (eye-tracking) with an offline measure (familiarity rating) to examine non-natives' processing of existing (congruent with the participants' L1, i.e., Arabic) and novel binomials during and after reading. Novel binomials were presented either once or six times in order to assess the effect of frequency on the learning of binomials. The aim was to address three research questions: (RQ1) Do non-native speakers demonstrate online sensitivity to the configuration (i.e., word order) of existing congruent binomials and novel binomials encountered in context?, (RQ2) How does the processing of novel binomials alter as a function of frequency of exposure (1 to 6 exposures)?, and (RQ3) Do non-native speakers exhibit offline sensitivity to the configuration of L1–L2 congruent existing binomials and novel binomials encountered in context?

https://doi.org/10.1017/S0142716426100538 Published online by Cambridge University Press

Regarding RQ1, we found that Arabic non-native speakers of English did not exhibit sensitivity to the preferred word order of existing binomials. This finding is in line with Sonbul et al.'s (2023) study which showed no online sensitivity to existing binomials for a similar group of non-native speakers. It should be noted, though, that in the present study, we only included existing binomials which are congruent in word order with Arabic. Thus, it seems that Arabic speakers of English are not sensitive to established configuration restrictions in binomials even when these are consistent with the restrictions in their L1. The results are also in line with Siyanova-Chanturia et al. (2011) in that non-natives are not always sensitive to the direction of established binomials. It should be noted that Siyanova-Chanturia et al. (2011) found a modulating effect of L2 proficiency on the processing of L2 binomials, with clearer sensitivity as proficiency increases, but the present study did not find such an effect. The following offers some potential explanations for this discrepancy. First, the participants in the two studies likely differed in their L2 English proficiency. Although there are no objective measures that would allow a direct comparison between non-native participants in Siyanova-Chanturia et al. (2011) and the present study, there are major differences between them. Participants in the former study were students at a university in the UK and were thus ESL (English-as-a-second-language) users who had spent an average of 21 months in an English-speaking community and had passed the requirement for studying at a British university (minimum 6.00 in IELTS). In contrast, our participants were EFL (English-as-a-foreign-language) speakers at the time of study (although some of them had spent some time in an English-speaking country). Yamashita and Jiang (2010) showed that EFL and ESL users process congruent/incongruent formulaic sequences (i.e., collocations) differently (with the effect of congruency diminishing for ESL users). Second, as discussed by Sonbul et al. (2023), the absence of a sensitivity in L1 Arabic speakers in L2 English might be related to the fact that binomials in Arabic are less fixed than those in English. Hence, if binomial word order is less fixed or biased in the L1, non-native speakers may be less attuned to this property in their L2. It would be interesting for future research to examine the effect of congruency on the processing of binomials by L1 Arabic-L2 English speakers in the EFL and ESL contexts.

For the novel binomials, the comparison between the first forward exposure and the reversed forms (for the 1-rep and 6-rep item types) consistently showed that the reversed form was processed faster than the forward form for the 6-rep items. This finding seems to suggest that when it comes to co-occurrence restrictions (e.g., *wires* used in combination with *pipes* regardless of the word order), seeing the binomials six times leads to a significant decrease in reading times. This finding is in line with previous research on non-natives' online processing of collocations (Altamimi & Conklin, 2024a; Durrant & Schmitt, 2010; Pellicer-Sánchez, 2017; Pellicer-Sánchez et al., 2022) and binomials (Sonbul et al., 2023) which showed that sensitivity to co-occurrence restrictions can be developed through repeated exposure in reading. This is also consistent with the findings of Sonbul et al. (2023) where more than one exposure (two or four) can lead to developing this co-occurrence restriction.

Turning now to RQ2, results comparing across the various occurrences of the 6-rep novel binomials (six forward and once reversed) showed no processing disadvantage for reversed forms in comparison to forward forms (as evident in Conklin &

https://doi.org/10.1017/S0142716426100538 Published online by Cambridge University Press

Carrol, 2021 for natives). On the contrary, more exposure to the novel binomials (including the reversed form) generally led to a decrease in reading times (see Figure 3). As for the comparison between the various forward forms and the reversed form, our findings suggest that it is only the first exposure that was processed significantly more slowly than the reversed form, again stressing the finding reported above that non-natives in the present study were sensitive to co-occurrence restrictions rather than word order preference. Hence, similar to Sonbul et al. (2023), it can be concluded that the Arabic-English non-native speakers treated the reversed form as another occurrence of the same binomial regardless of word order. This is the case despite the fact that the present study added two more exposures to novel binomials (i.e., the maximum number of exposures was four in Sonbul et al. 2023 but six in the present study). It seems that increasing exposure did not alter the tendency for non-natives to be unattuned to word order when processing L2 novel binomials. This is in contrast with the native speakers in Conklin and Carrol's (2021) study who showed a processing advantage for novel binomials presented in the forward form over their reversed forms after four to five exposures. Thus, non-natives in the present study appear to perform differently than natives in that they do not seem to be sensitive to configuration restrictions when processing binomials online, possibly due to binomials being not as fixed in Arabic as they are in English (see above).

In addition to the online (eye-tracking) measure, the present study also included a post-exposure offline rating task to examine declarative knowledge of existing binomials (RQ3) and novel binomials after exposure (RQ4). For existing binomials which are congruent with the speakers' L1 (Arabic), results showed higher ratings for the forward form over the reversed form for existing binomials (i.e., higher ratings) regardless of whether these were presented during the reading task or not. Thus, similar to Sonbul (2015) and Sonbul and El-Dakhs (2020) who showed dissociation between offline and online processing of collocations, the current study seems to suggest a similar dissociation between online (eye-movements) and offline (familiarity rating) processing of existing binomials. While offline, non-natives exhibited declarative knowledge of configuration restrictions congruent with their L1, they did not show such sensitivity online. Thus, declarative knowledge of formulaic sequences might develop earlier than online sensitivity which may require more exposure to be apparent (Sonbul, 2015).

Turning to the findings of the familiarity rating task in relation to novel binomials, non-native speakers in the present study did not show sensitivity to configuration restrictions offline; the novel binomials were rated similarly in both directions. This finding stands in contrast with several previous studies such as Sonbul and Schmitt (2013), Altamimi and Conklin (2024b) and Alotaibi et al. (2022). For example, in a study examining collocations, Sonbul and Schmitt (2013) found that non-native speakers developed declarative knowledge (recall and recognition) of unfamiliar, medical collocations after three exposures even though they did not manifest any online sensitivity. Similarly, previous research on binomials showed evidence for non-natives' developing offline sensitivity to configuration restrictions after two exposures (Altamimi & Conklin, 2024b) and six exposures (Alotaibi et al., 2022). It should be noted though that these previous studies on binomials used different forms of declarative measures than the present study. While in the present study, the items were presented in both the forward and

https://doi.org/10.1017/S0142716426100538 Published online by Cambridge University Press

reversed forms in the familiarity rating task to allow a direct comparison, Altamimi and Conklin (2024b) used a forced-choice measure (Is it *gold and silver* or *silver and gold*) and Alotaibi et al. (2022) used a multiple-choice test and a familiarity rating task in the forward direction only. We argue that the offline measure employed in the present study might be more appropriate for detecting sensitivity to configuration restrictions through the direct comparison of the forward and reversed forms of the same binomial (although see below for a limitation in our offline task).

Overall, the results of the present study suggest that Arabic speakers of English are not sensitive to the preferred word order (i.e., configuration) for either existing or novel binomials online. Offline, they seem to be sensitive to binomials' configuration restrictions for items that exist both in their L1 and L2 with the same preferred configuration (e.g., *salt and pepper* over *pepper and salt*). However, they do not develop such sensitivity to novel binomials after encountering them six times in a reading context.

Limitations and suggestions for future research

The study is limited in several ways. First, we only included existing binomials that have the same word order preference in L1 and L2 (i.e., congruent binomials). Future research on the processing of binomials could include both congruent and incongruent binomials and involve both offline and online measures. This would allow a more direct investigation of the effect of congruency on the processing of L2 binomials whether in real time during reading (eye-tracking) or after reading (declarative post-exposure tests).

Additionally, the offline measure employed in the present study was limited in that it presented both the forward and reversed forms to the participants without any counterbalancing (though in a balanced design), which might have resulted in them developing strategies in rating the binomials (the first presentation of the binomial—whether forward or reversed—might have an influence on the second one). Future research can attempt to develop counter-balanced lists to avoid such task-order effects on findings.

Third, and most importantly, the current research is limited in that it only included six exposures to the target novel binomials, which, based on the results, was not sufficient for the non-natives to develop an offline or online sensitivity to their configuration. Further research in this area will benefit from adding extra exposures to novel binomials in context. For example, Pellicer-Sánchez et al. (2022) found that the processing of novel collocations became similar to existing ones only after eight exposures. Thus, presenting novel binomials eight or even 10 times might help address the issue of how many exposures are needed before non-natives can develop sensitivity to novel binomials. Developing sensitivity to configuration restrictions might also pertain to longer-term, more consistent exposure, rather than several encounters in a short period of time. Hence, future research can adopt a longitudinal design where participants see the target binomial repeatedly in different contexts over a long period of time.

Another area for future research to consider is the inclusion of an explicit instruction or textual enhancement component instead of presenting binomials in

https://doi.org/10.1017/S0142716426100538 Published online by Cambridge University Press

normal font as was the case in most previous research on L2 binomials (Alotaibi et al., 2022; Altamimi & Conklin, 2024b; Sonbul et al., 2023). Previous research on the acquisition of collocational co-occurrence restrictions showed an advantage for textual enhancement and explicit instruction over natural exposure on offline measures (e.g., Sonbul & Schmitt, 2013). Thus, future research can possibly examine the effect of such exposure conditions on the ability of non-natives to develop sensitivity to the configuration restrictions of novel binomials. This would have tangible implications for the teaching of binomials in L2 classes.

Finally, as indicated above, Arabic is less fixed than English when it comes to binomial configuration restrictions. Further research in the area can attempt to compare across non-native participants from various L1s—more or less fixed in the word order of binomial components. This would help address the issue of whether the current results are due to a transfer effect whereby certain L1s have less rigid word order preferences than others, and this is transferred into L2 processing. The question then is are non-natives better at watching out for L2 restrictions (e.g., "salt and pepper" being more acceptable than "pepper and salt") to which they are attuned in their L1?

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0142716426100538

Replication package. Replication data and materials for this article can be found at https://osf.io/x4whb.

Acknowledgements. The researchers thank Prince Sultan University in Saudi Arabia for funding this research project through the grant SEED-CHS-2023-130.

Competing interests. The author(s) declare none.

Notes

1 This ratio follows Mollin's (2012) classification of binomials based on frequency of occurrence. She sets a threshold of 75% (i.e., 3:1 ratio) for moderate preference and 90% (i.e., 9:1 ratio) for strong preference. Thus, all our existing binomials have moderate to strong word order preference both in Arabic and English.

2 The only exception was the item 'life and death' (existing control item) with a ratio of 1:1.78. This item was included in the pool as, according to the judgment of two native speakers of Arabic, 'life and death' was more common than 'death and life'. We also re-ran the analysis for offline data (RQ3) excluding this item, and the results were identical to those reported in the manuscript.

3 There were a few exceptions (*spoons and bowls, ability and strength, tennis and football, grass and leaves*) where the ratio in English or Arabic exceeded 2, but since the frequency in both directions was low, we kept these items in the pool.

4 It should be noted that since the effects at Word 1 are not relevant to the present study (as we do not expect any sensitivity to emerge before the participant sees the last word of the binomial), we focus here only on Word 3 and the whole phrase. Results related to Word 1 analysis are available online via the same OSF link referenced above.

5 We ran all eye-tracking analyses with the full data set, and the results were the same as the ones reported here.

6 The whole phrase frequency factor significantly improved the model fit for the TRT measure (both Word 3 and whole phrase). However, since the VIF values showed a collinearity issue with VIF values exceeding 10, we excluded this factor from the final model.

https://doi.org/10.1017/S0142716426100538 Published online by Cambridge University Press

References

Alotaibi, S., Pellicer-Sánchez, A., & Conklin, K. (2022). The effect of input modes and number of exposures on the learning of L2 binomials. *ITL – International Journal of Applied Linguistics*, **173**, 58–93.

Arai, Y., & Takizawa, K. (2024). Text and reading task variables in incidental L2 vocabulary learning from reading: A methodological synthesis. *Research Methods in Applied Linguistics*, **3**(2). https://doi.org/10.1016/j.rmal.2024.100110

Altamimi, A. (2021). The Acquisition and Processing of Formulaic Language: A Focus on Binomials and Collocations and the Effects of Congruency, Frequency and Learning Conditions. PhD thesis, University of Nottingham.

Altamimi, A., & Conklin, K. (2024a). Effects of pre-reading study and reading exposure on the learning and processing of collocations. *TESOL Quarterly*, **58**, 1324–1346. https://doi.org/10.1002/tesq.3268

Altamimi, A., & Conklin, K. (2024b). The effect of congruency and frequency of exposures on the learning of L2 binomials. *Languages*, **9**(1), 9. https://doi.org/10.3390/languages9010009

Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language*, **62**, 67–82.

Arts, T., Belinkov, Y., Habash, N., Kilgarriff, A., & Suchomel, V. (2014). arTenTen: Arabic corpus and word sketches. *Journal of King Saud University-Computer and Information Sciences*, **26**(4), 357–371.

Benor, S., & Levy, R. (2006). The chicken or the egg? A probabilistic analysis of English binomials. *Language*, **82**, 233–278. https://doi.org/10.1353/lan.2006.0077

Brysbaert, M., Mandera, P., & Keuleers, E. (2018). The word frequency effect in word processing: An updated review. *Current Directions in Psychological Science*, **27**(1), 45–50.

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.

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.

Conklin, K., & Carrol, G. (2019). First language influence on the processing of formulaic language in a second language. In A. Siyanova-Chanturia & A. Pellicer-Sánchez (Eds.), *Understanding formulaic language: A second language acquisition perspective* (pp. 62–77). Routledge.

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.

Conklin, K., & Carrol G. (2021). Words go together like 'bread and butter': The rapid, automatic acquisition of lexical patterns. *Applied Linguistics*, **42**(3), 492–513

Davies, M. (2016) *Corpus of News on the Web (NOW)*. Available online at https://www.english-corpora.org/now/.

Durrant, P., & Schmitt, N. (2010). Adult learners' retention of collocations from exposure. *Second Language Research*, **26**(2), 163–188. https://doi.org/10.1177/0267658309349431

El-Dakhs, D. A. S., Sonbul, S., & Masrai, A. (2024). An eye-tracking study on the processing of L2 collocations: The effect of congruency, proficiency, and transparency. *Journal of Psycholinguistic Research*, **53**(2), 30.

Elgort, I., Brysbaert, M., Stevens, M., & Van Assche, E. (2018). Contextual word learning during reading in a second language: An eye-movement study. *Studies in Second Language Acquisition*, **40**(2), 341–366.

Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text-Interdisciplinary Journal for the Study of Discourse*, **20**(1), 29–62.

Godfroid, A., Ahn, J., Choi, I., Ballard, L., Cui, Y., Johnston, S., Lee, S., Sarkar, A., & Yoon, H. J. (2018). Incidental vocabulary learning in a natural reading context: An eye-tracking study. *Bilingualism: Language and Cognition*, **21**, 563–584.

Godfroid, A., Boers, F., & Housen, A. (2013). An eye for words: Gauging the role of attention in incidental L2 vocabulary acquisition by means of eye-tracking. *Studies in Second Language Acquisition*, **35**(3), 483–517.

Gorgis, D. T., & Al-Tamimi, Y. (2005). Binomials in Iraqi and Jordanian Arabic. *Journal of Language and Linguistics*, **4**(2), 135–151.

Joseph, H. S. S. L., Wonnacott, E., Forber, P., & Nation, K. (2014). Becoming a written word: Eye movements reveal order of acquisition effects following incidental exposure to new words during silent reading. *Cognition*, **133**, 238–248.

https://doi.org/10.1017/S0142716426100538 Published online by Cambridge University Press

Kiss, G. R., Armstrong, C., Milroy, R., & Piper, J. (1973). An associative thesaurus of English and its computer analysis. In A. J. Aitken, R. W. Bailey, & N. Hamilton-Smith (Eds.), *The computer and literary studies* (pp. 153–65). Edinburgh University Press.

Laufer, B., & Rozovski-Roitblat, B. (2011). Incidental vocabulary acquisition: The effects of task type, word occurrence and their combination. *Language Teaching Research*, **15**(4), 391–411.

Majuddin, E., Siyanova-Chanturia, A., & Boers, F. (2022). Incidental acquisition of multiword expressions through audiovisual materials: The role of repetition and typographic enhancement. *Studies in Second Language Acquisition*, **43**(5), 985–1008.

Meara, P., & Miralpeix, I. (2017). *Tools for researching vocabulary*. Multilingual Matters.

Mohamed, A. A. (2018). Exposure frequency in L2 reading: An eye-movement perspective of incidental vocabulary learning. *Studies in Second Language Acquisition*, **40**, 269–293. https://doi.org/10.1017/s0272263117000092

Mollin, S. (2012). Revisiting binomial order in English: ordering constraints and reversibility. *English Language & Linguistics*, **16**(1), 81–103.

Morgan, E., & Levy, R. (2016). Abstract knowledge versus direct experience in processing of binomial expressions. *Cognition*, **157**, 384–402.

Nation, I. S. P. (2012). The BNC/COCA word family lists. Document bundled with Range Program with BNC/COCA Lists, 25. Retrieved from: https://www.victoria.ac.nz/lals/about/staff/publications/paul-nation/Information-on-the-BNC_COCA-word-family-lists.pdf

Pellicer-Sánchez, A. (2016). Incidental L2 vocabulary acquisition from and while reading: An eye-tracking study. *Studies in Second Language Acquisition*, **38**(1), 97–130.

Pellicer-Sánchez, A. (2017). Incidental learning of L2 collocations: A classroom study. *Language Teaching Research*, **21**(3), 381–402. https://doi.org/10.1177/1362168815618428

Pellicer-Sánchez, A., & Boers, F. (2019). Pedagogical perspectives to the teaching and learning of formulaic language. In A. Siyanova-Chanturia & A. Pellicer-Sánchez (Eds.), *Understanding formulaic language: A second language acquisition perspective* (pp. 153–173). Routledge.

Pellicer-Sánchez, A., & Schmitt, N. (2010). Incidental vocabulary acquisition from an authentic novel: Do" things fall apart"? *Reading in a Foreign Language*, **22**(1), 31–55.

Pellicer-Sánchez, A., Siyanova-Chanturia, A., & Parente, F. (2022). The effect of frequency of exposure on the processing and learning of collocations: A comparison of first and second language readers' eye movements. *Applied Psycholinguistics*, **43**(3), 727–756.

Pitts, M., White, H., &Krashen, S. (1989). Acquiring the second language vocabulary through reading: A replication of the Clockwork Orange study using second language acquirers. *Reading in a Foreign Language*, **5**, 271–275.

R Core Team (2024). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available online at https://www.R-project.org/.

Siyanova-Chanturia, A., Conklin, K., & Van Heuven, W. J. (2011). Seeing a phrase "time and again" matters: The role of phrasal frequency in the processing of multiword sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **37**(3), 776.

Siyanova-Chanturia, A., Conklin, K., Caffarra, S., Kaan, E., & van Heuven, W. J. (2017). Representation and processing of multi-word expressions in the brain. *Brain and Language*, **175**, 111–122.

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.

Sonbul, S., & El-Dakhs, D. (2020). Timed versus untimed recognition of L2 collocations: Does estimated proficiency modulate congruency effects? *Applied Psycholinguistics*. **41**, 1197–1222. https://doi.org/10.1017/S0142716420000135

Sonbul, S., & Schmitt, N. (2013). Explicit and implicit lexical knowledge: Acquisition of collocations under different input conditions. *Language Learning*, **63**(1), 121–159. https://doi.org/10.1111/j.1467-9922.2012.00730.x

Sonbul, S., El-Dakhs, D. A. S., Conklin, K., & Carrol, G. (2023). "Bread and butter" or "butter and bread"? Nonnatives' processing of novel lexical patterns in context. *Studies in Second Language Acquisition*, **45**(2), 370–392.

Tremblay, A., Derwing, B., Libben, G., & Westbury, C. (2011). Processing advantages of lexical bundles: Evidence from self-paced reading and sentence recall tasks. *Language Learning*, **61**(2), 569–613.

https://doi.org/10.1017/S0142716426100538 Published online by Cambridge University Press

Van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology*, **67**(6), 1176–1190.

von der Malsburg T., & Angele, B. (2017). False positives and other statistical errors in standard analyses of eye movements in reading. *Journal of Memory and Language*, **94**, 119–133.

Waring, R., & Takaki, M.(1989). At what rate do learners learn and retain new vocabu lary from reading a graded reader?. *Reading in a Foreign Language*, **15**, 130–163.https://doi.org/10.64152/10125/66776

Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied linguistics*, **28**(1), 46–65.

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

Yamashita, J., & Jiang, N. A. N. (2010). L1 influence on the acquisition of L2 collocations: Japanese ESL users and EFL learners acquiring English collocations. *TESOL Quarterly*, **44**(4), 647–668. https://doi.org/10.5054/tq.2010.235998

Cite this article: Sonbul, S., Carrol, G., El-Dakhs, D. A. S., Conklin, K., & Altamimi, A. (2026). Learning binomials from context: Do non-natives develop offline and online sensitivity to novel binomials after exposure? *Applied Psycholinguistics*. https://doi.org/10.1017/S0142716426100538

https://doi.org/10.1017/S0142716426100538 Published online by Cambridge University Press