LANGUAGE LEARNING
A Journal of Research in Language Studies
Language Learning ISSN 0023-8333

EMPIRICAL STUDY

Collocation in the Mind: Investigating Collocational Priming in Second Language Speakers of Italian

Irene Fioravanti,[a] Anna Siyanova-Chanturia,[b] and Alessandro Lenci[c]

[a]University for Foreigners of Perugia [b]Te Herenga Waka – Victoria University of Wellington

[c]University of Pisa

Abstract: Collocational priming is a priming effect induced by collocationally related words; it has been taken to explain the cognitive reality of collocation. Collocational priming has largely been observed in first language (L1) speakers, whereas work on the representation of collocation in a second language (L2) is still limited. In the present study, we sought to investigate this phenomenon in L1 and L2 speakers of Italian. We used a lexical decision task to explore collocational priming in verb–noun and noun–adjective collocations differing in frequency and collocational strength. Both L1 and L2 speakers were found sensitive to the frequency of collocations. Importantly, exposure to L2 Italian was found to play a role. The results suggest that collocational priming occurs both in L1 and L2 speakers, and that the mechanisms associated with collocation processing and representation in L1 and L2 speakers may be comparable.

Keywords collocation; priming; phrase frequency; mental lexicon; second language; Italian

---

CRediT author statement – **Irene Fioravanti**: conceptualization; methodology; investigation; data analysis; writing – original draft preparation; writing – review and editing. **Anna Siyanova-Chanturia**: conceptualization; methodology; writing – review and editing. **Alessandro Lenci**: conceptualization; writing – review and editing.

A one-page Accessible Summary of this article in nontechnical language is freely available in the Supporting Information online and at https://oasis-database.org

Correspondence concerning this article should be addressed to Dr. Irene Fioravanti, University for Foreigners of Perugia, Piazza Fortebraccio 4, 06123 Perugia, Italy. Email: irene.fioravanti@unistrapg.it

The handling editor for this manuscript was Kristopher Kyle.

© 2024 Language Learning Research Club, University of Michigan.
DOI: 10.1111/lang.12663

Introduction

Recent years have seen a growing interest in the online processing of collocation (for a review, see Siyanova-Chanturia & van Lancker-Sidtis, 2019). This is largely due to the observation that collocations are pervasive in language (Hoey, 2005), and are characterized by "mutual expectancy" (Firth, 1957), such that where one collocation constituent is found, the other one is likely to be found next to it. Hoey (1991) operationalized Firth's notion of "mutual expectancy" as "the relationship a lexical item has with items that appear with greater than random probability in its textual context" (pp. 6–7). However, Hoey (2005) moved beyond a textual definition of collocations (Evert, 2008; Firth, 1957; Sinclair, 1991), proposing a model that links collocations in text to collocations in the mind. Hoey (2005) defined a collocation as "a psychological association between words (rather than) lemmas up to four words apart"; this psychological association "is evidenced by their occurrence together in corpora more often than is explicable in terms of random distribution" (p. 5). This definition was intended to posit collocation as "a psycholinguistic phenomenon, the evidence for which can be found statistically in computer corpora" (Hoey, 2005, p. 5). As Hoey argued, this psychological association can be explained in terms of the psycholinguistic phenomenon of priming (Meyer & Schvaneveldt, 1971), a mechanism whereby a word (e.g., *doctor*) is recognized faster when it is preceded by a related word (e.g., *nurse*) than when it is preceded by an unrelated one (e.g., *butter*).

Drawing on the concept of priming, Hoey (2005) put forward the idea of collocational priming, which highlights the link between collocation in text and collocation in the mind. Collocational priming is a priming effect between the constituents of a collocation. As Hoey (2005) noted, priming mechanisms are at the root of language learning. Learners acquire a word when they encounter it in text or spoken discourse and derive information about the context and the cotext in which the word occurs. Speakers store information about the tendency of words to co-occur with other words in the mental lexicon. At the same time, Hoey (2005) argued that collocations are produced in a predictable way because of the priming relationships between their constituents. For example, when a language user encounters the first element of a collocation (e.g., *heavy*), the recall of the second element (e.g., *rain*) is facilitated, due to the existence of the frequent phrase *heavy rain*. Furthermore, Hoey (2005) claimed that every time a speaker uses a word, its use reinforces the existing relationship between that word and the context in which it occurs; or it may weaken it if the word is encountered in an unfamiliar context. As Hoey (2005) argued, a speaker's mind creates a mental concordance of every word it has encountered.

This mental concordance can be easily accessed, such that all kinds of patterns and configurations, including collocational patterns, are available for use. Importantly, according to Hoey (2005), there may be little difference in the collocational priming processes between first language (L1) and second language (L2) speakers; rather, the key difference lies in the amount of exposure to the language. If learners are sufficiently exposed to a L2, the priming mechanisms at the root of the acquisition of L2 collocations may be comparable to the priming mechanisms responsible for the acquisition of L1 collocations.

Hoey's (2005) theory provides a model of how frequency may relate to a linguistic item's mental representation, suggesting that higher frequency collocates are more mentally primed with each other than lower frequency ones. Thus, Hoey's (2005) theory of collocational priming has been taken to support the cognitive reality of collocations. However, as Durrant and Doherty (2010, p. 126) pointed out, directly linking corpus data with cognitive data is problematic. They argue that it is unlikely that any corpus can match the linguistic experience of a given speaker, since its content depends on the purpose for which it was built. Large corpora may cover a range of discourse types; in contrast, small corpora may be representative of one particular variety of language or genre. Therefore, one might expect differences between a corpus and a language user's experience, raising the question of whether corpus-based collocations have a cognitive correlate in the language user's mind (Durrant & Doherty, 2010). Accordingly, one should be cautious when using corpus evidence to support claims about the human mind. As suggested by Durrant and Siyanova-Chanturia (2015), corpus data can help researchers to formulate hypotheses specific to language processing, but it is important that these hypotheses are also empirically tested. As Hoey (2005) noted, a corpus is likely to be representative of the type of input a speaker would have encountered and can be considered as a valid tool in the investigation of the psycholinguistic reality of collocations. We thus believe there is much to be gained from integrating corpus data with psycholinguistic approaches in examining the cognitive reality of collocations.

The overarching aim of the present study is to test Hoey's (2005) theory of collocational priming by examining whether collocations that vary in corpus-based measures (such as frequency and association strength) have psychological reality for L1 and L2 speakers of Italian.

Background Literature

One of the first studies to investigate whether corpus-based measures—namely, frequency and mutual information (MI)—affect L1 and L2 processing of common phrases was that of Ellis et al. (2008; Study 3). MI is an association measure that shows how strongly associated the words are, based on their probability to occur together (Evert, 2008). Ellis et al. (2008) employed a naming task to investigate whether the articulation of the last word within a sequence (i.e., *words*) was affected by the earlier part of it (i.e., *in other*). Their results showed that MI significantly affected the articulation of the last word by L1, but not L2, speakers. However, the articulations of L2, but not L1, speakers were affected by frequency. According to Ellis et al. (2008), L1 speakers' sensitivity to MI suggests that these participants perceived common sequences as holistic units. In contrast, L2 learners perceived these sequences as high-frequency items, but not as prefabricated, holistic patterns. In addition, using a lexical decision task (LDT), Ellis et al. (2009; Experiment 1) investigated whether lexical access was affected by the frequency of collocations in L1 speakers of English. The results showed a significant priming effect between the elements of high-frequency collocations (e.g., *end* and *war*), whereas no priming effect was found between the elements of low-frequency collocations (e.g., *begin* and *afresh*). However, one should be cautious in interpreting the results of Ellis and colleagues (2008, Study 3; 2009, Experiment 1) due to the low number of participants (i.e., 18 L1 vs. 16 L2 speakers, and 15 L1 speakers, respectively).

A clearer picture of the collocational priming phenomenon comes from Durrant and Doherty (2010). These authors carried out two LDTs to investigate whether a priming effect can be found for frequent collocations, associated frequent collocations, or both, in L1 speakers of English. Associated and non-associated collocations were used as two different experimental conditions. Of note is that collocation elements are judged by speakers to be associates only if the second element is listed as an associate of the first one in free association norms, in which participants are asked to produce the first word that comes to mind when they read the stimulus word. In contrast, if the second element is not listed as an associate of the first one, collocation elements are judged to be non-associated. In a study by Durrant and Doherty (2010), the association strength of collocation elements was assessed via the Edinburgh Associative Thesaurus and in a word association task with L1 speakers of English. Collocations belonged to one of four experimental conditions: associated frequent collocations (e.g., *card game*), non-associated frequent collocations (e.g., *mental picture*), non-associated moderate-frequency collocations (e.g., *greater concern*), and non-associated low-frequency collocations (e.g., *weak ground*). In the first LDT, participants responded faster to non-associated frequent collocations and to associated frequent collocations than to their respective controls (i.e., combinations which did not occur in the corpus; these items were created by substituting the first element of the collocation, e.g., *waiting room > short room*). No priming was found in the low-frequency and moderate-frequency conditions. However, the intervention of strategic processes (i.e., participants being able to find a relation between the prime and the target) could not be ruled out. To minimize the likelihood of strategic processes (i.e., participants anticipating the target word), a pattern mask is typically used to elicit an automatic priming. The pattern mask is presented before the prime as a set of hashtags or asterisks (Jiang, 2012). Thus, in the second LDT, a pattern mask was used by Durrant and Doherty. Results showed a significant priming effect only between associated frequent collocations, with no priming effect found for frequent collocations. Findings suggest that the priming between associates may be controlled by automatic processes, whereas the priming between collocations may be governed by strategic processes.

The study by Durrant and Doherty (2010) was more recently extended by Cangır et al. (2017), who investigated collocational priming in L1 speakers of Turkish. Statistical and semantic criteria were used to extract verb–noun (V + N) and adjective–noun (Adj + N) collocations from corpora. A LDT was used with a pattern mask to ensure automatic priming. Results provided evidence of collocational priming in Turkish, with phrase frequency being a significant predictor of reaction times (RTs). In addition, the frequency of the second word of a collocation was found significant, suggesting that both phrase frequency and word frequency may affect collocation processing.

The studies reviewed above have largely focused on collocational priming in L1 speakers (except for Ellis et al., 2008). Although considerable work has recently been done on collocation processing in L2 speakers, much of this has been concerned with the effect of frequency and/or congruency (i.e., having an equivalent L1 construction) on the processing of L2 collocations (e.g., Du et al., 2023; Gyllstad & Wolter, 2016; Sonbul & El-Dakhs, 2020; Wolter & Gyllstad, 2013). Research into L2 collocation processing has lent some support to Hoey's (2005) model of collocational priming, suggesting that higher frequency L2 collocations are processed faster than lower frequency collocations (Sonbul, 2015). Psycholinguistic studies have also demonstrated that L2 speakers are sensitive to corpus-derived association measures (e.g., Öksüz et al., 2020). To the best of our knowledge, only three published studies have so far used the priming paradigm to test the collocational priming theory in a L2. First, Sonbul and Schmitt (2013) focused on the relation between implicit and explicit knowledge of collocations. They used low-frequency medical collocations (e.g., *vanishing lung*), as such items were very likely to be unknown to the participants. L2 proficiency was assessed through a self-assessment questionnaire targeting different language skills. L1 and L2 speakers of English were exposed to three different learning conditions varying in terms of contextual exposure (enriched, enhanced, and decontextualized); this allowed the researchers to observe whether the type of learning condition affected the development of implicit knowledge (tested with a priming task). Results showed that neither of the two groups of speakers developed implicit collocational knowledge. Sonbul and Schmitt took their results to support Hoey's (2005) tenet that large amounts of input are necessary to acquire collocations in both a L1 and a L2.

Second, Toomer and Elgort (2019) replicated and extended the results of Sonbul and Schmitt (2013) with advanced learners of English. Lexical proficiency was measured through the Vocabulary Size Test (Nation & Beglar, 2007). The original design was modified by increasing the number of contextual exposures and extending the learning sessions from one to two days. Toomer and Elgort used a primed LDT in order to test the development of implicit knowledge. The second element of collocations was used as the target (e.g., *baby*), preceded by the prime (e.g., *cloud*) or another word that was not a collocate of the target (e.g., *steam*). Contrary to Sonbul and Schmitt's (2013) findings, Toomer and Elgort found evidence for the development of implicit knowledge of medical collocations. They argued that "the priming effect…was likely due to a larger number of exposures to the collocations" (Toomer & Elgort, 2019, p. 19).

Third, crosslinguistic priming was investigated between L1 Turkish and L2 English by Cangır and Durrant (2021). Proficiency was self-reported by learners in a language background questionnaire, and vocabulary knowledge was assessed with Nation and Beglar's (2007) Vocabulary Size Test. Cangır and Durrant investigated whether crosslinguistic priming was affected by the part of speech of the target items (Adj + N vs. V + N collocations), L1–L2 congruency, the presentation direction (L1–L2 vs. L2–L1), or phrase frequency. A masked LDT was used to elicit automatic priming. Evidence for crosslinguistic priming was found in Adj + N collocations, but not in V + N collocations. According to Cangır and Durrant, this was due to the fact that Adj + N collocations have the same word order in Turkish and English, whereas the word order of V + N collocations differs between the two languages. Participants responded faster to Adj + N collocations in the L1–L2 order than in the L2–L1 order. Furthermore, L1 adjective primes facilitated the recognition of English noun targets where the collocations were congruent between the two languages. Phrase frequency was found to affect participants' RTs: High-frequency collocations received faster responses than low-frequency collocations. A key contribution of this study by Cangır and Durrant is that it extended Hoey's (2005) work by investigating the theory of collocational priming from a crosslinguistic perspective. The results suggest that the mechanism of priming in a L1 and a L2 might be comparable, in that the activation of individual words in the L1 spreads to the collocates of those words in the L2. Further, word order may sometimes play a helpful role in strengthening the associative links between collocation elements, in that collocations with the same word order in the two target languages may be more likely to be entrenched in memory than collocations with a different word order.

The Present Study

The findings of the above studies strongly suggest that frequency and corpus-derived association measures affect collocation processing in both a L1 (Cangır et al., 2017; Durrant & Doherty, 2010) and a L2 (Cangır & Durrant, 2021; Toomer & Elgort, 2019). Moreover, repeated exposure to the input may be helpful in learning collocations as it strengthens the associative links between collocation components (Sonbul & Schmitt, 2013; Toomer & Elgort, 2019). Finally, the findings lend support to Hoey's (2005) claim that the mechanisms at the root of collocational priming in L1 and L2 speakers may be comparable. However, several gaps remain to be addressed. First, the question of whether or not collocations from different frequency bands elicit a priming effect also in L2 speakers has so far not been investigated. Second, the amount of exposure to the L2 has not been considered as a predictor in previous studies. Third, although L2 proficiency has been recognized as an important variable in L2 online processing, it is not clear whether proficiency modulates collocational priming in a L2. In order to address these gaps, we sought to answer the following research questions:

1. Does collocational priming occur in L1 and L2 speakers of Italian?

If the answer to the first question is affirmative, the role of other variables at the root of collocational priming will be investigated with the following questions:

2. Does phrase frequency modulate the strength of collocational priming?
3. Do exposure to the L2 and L2 proficiency play a role in strengthening the associative links between collocation constituents?
4. To what extent are the mechanisms at the root of L2 collocational priming comparable to those responsible for L1 collocational priming?

Based on existing literature, we hypothesized that collocational priming occurs in both L1 and L2 speakers of Italian. To test this hypothesis, we employed a LDT with L1 and L2 speakers of Italian, using a corpus-based approach to defining and identifying collocations. As our study investigated whether the first word of a collocation facilitated the recognition of the second word, L1 and L2 speakers were presented with the first term of a collocation and were asked to make lexical decisions on the second word. The RTs on the second word of the collocation were compared with those on the second word of the control pair.

Further, as previous studies have showed that frequency has an impact on the elicitation of collocational priming (Cangır et al., 2017; Cangır & Durrant, 2021; Durrant & Doherty, 2010), the items in the present study were selected from different frequency bands. This design allowed us to investigate the phenomenon of collocational priming in three experimental conditions: high frequency, medium frequency, and low frequency. Based on the results of previous studies, we hypothesized that collocational priming would be found only in the high-frequency condition. As Hoey (2005) suggested, a repeated exposure to the input may be helpful in strengthening the associative links between collocation components. Therefore, L2 speakers sufficiently exposed to the target language might show similar patterns to L1 speakers in the processing of collocations. In contrast, little exposure to L2 and little experience with the target language will not strengthen the priming mechanisms at the root of the recognition of collocations during online processing, and thus collocation processing in a L2 may differ from that in a L1. Therefore, we predicted that high-proficiency learners with a greater amount of exposure to the L2 should respond faster to collocations compared to control pairs. By contrast, we did not expect a priming effect in lower proficiency learners with less exposure to Italian. Finally, Hoey (2005) argued that there may be little difference in collocational priming processes between L1 and L2 speakers, as the key difference may lie in the amount of exposure to the language. We thus predicted that the mechanisms at the root of L2 collocational priming should be comparable to those responsible for L1 collocational priming.

Method

Participants

In total, 38 L1 speakers (30 females; age range: 19–35 years, $M = 25.1$, $SD = 4.4$) and 32 L2 speakers of Italian (23 females; age range: 19–45 years, $M = 25.3$, $SD = 6.5$) took part in the study. L1 and L2 participants were undergraduate and postgraduate students at one university in New Zealand and one in

Table 1 Learners' self-reported experience of Italian as a second language

Variables	*M*	*SD*	Range
First contact with Italian (in years)	20.5	6.5	9–40
Exposure to Italian (in months)[a]	2.8	1.1	1–4
Speaking[b]	3.3	0.9	1–5
Writing[b]	3.3	0.7	1–5
Listening[b]	3.9	0.9	2–5
Reading[b]	3.9	0.7	2–5

Note. [a]Based on a 4-point scale (1 = *never been*, 2 = *6 months or less*, 3 = *12 months or less*, 4 = *more than 12 months*). [b]Based on a 5-point scale (1 = *very poor*, 2 = *weak*, 3 = *ok*, 4 = *very good*, 5 = *excellent*).

Italy. L2 speakers of Italian came from the following L1 backgrounds: Arabic (2), English (11), Filipino (1), Persian (1), Polish (2), Portuguese (1), Russian (2), Slovakian (1), Slovenian (1), Spanish (5), German (4), and Turkish (1). Learners completed a language background questionnaire about their prior experience with the Italian language (Table 1). They further rated their speaking, writing, listening, and reading on a 5-point Likert scale (1 = *very poor*, 2 = *weak*, 3 = *ok*, 4 = *good*, 5 = *excellent*) and specified their level according to the Common European Framework of Reference for Languages (CEFR; i.e., A1, A2, B1, B2, C1, or C2). Learners in Italy were students enrolled in an Italian language program at a university and were living in Italy at the time of data collection. They were assigned to each class based on their CEFR level, which was assessed before the beginning of the Italian course. Learners in New Zealand were all undergraduate students of Italian.

Materials

Extraction of Target Collocations

We used Italian V + N and N + Adj collocations.[1] In V + N collocations, verbs were used as primes and nouns as target words (e.g., *mantenere* 'keep' prime – *promessa* 'promise' target). In N + Adj collocations, nouns were used as primes and the adjectives as target words (e.g., *agenda* 'diary/schedule' prime – *piena* 'full' target). Italian collocations were extracted from the *La Repubblica* corpus (Baroni et al., 2004), which is one of the largest corpora of the Italian language, comprising about 380 million words drawn from the daily newspaper *La Repubblica*. The corpus is tagged for part of speech and is lemmatized (Baroni et al., 2004). The accuracy of the tagging reached 97%. We extracted both types of collocations from the reference corpus using LexIt

Table 2 Frequency and mutual information bands of collocations

	Frequency			Mutual information		
Band	*M*	*SD*	Range	*M*	*SD*	Range
High	816.2	315.0	444–1,668	8.4	1.4	6.52–11.45
Medium	200.8	85.4	100–372	5.9	0.7	4.34–6.84
Low	49.1	23.1	19–95	4.3	0.5	3.01–4.93

(Lenci et al., 2012), a computational framework for the exploration of distributional properties of verbs, nouns, and adjectives in Italian. For each verb and noun, we explored their syntactic profile in the following syntactic slots: *#obj* (i.e., direct object of the verb) and *#modadj* (i.e., the adjective-modifier of the noun). Using LexIt, we preliminarily extracted 80 V + N collocations and 80 N + Adj collocations from the *La Repubblica* corpus.

Classification of Collocations by Frequency Band

We identified collocations using a frequency- and association-based approach (Evert, 2008). For each item, absolute frequency was derived from the *La Repubblica* corpus. Further, we calculated MI to identify collocations. Candidate items were identified as collocations if they had an MI score of 3 or above. The MI value of 3 or above is traditionally used as a threshold to identify "true" collocations (Evert, 2008). It is important to note, however, that this threshold is arbitrary and as such it has been criticized (e.g., Eguchi & Kyle, 2023; Gablasova et al., 2017). In addition, we used three bands of phrase frequency in order to investigate whether higher frequency word pairs were more likely to elicit collocational priming than lower frequency ones. We logarithmically transformed phrase frequency and assigned each candidate item to one of the following frequency bands: high frequency, medium frequency, or low frequency. From the preliminary pool of 160 collocations, we then sorted collocations by phrase frequency and selected 42 V + N collocations (e.g., *mantenere una promessa* "to keep a promise") and 42 N + Adj collocations (e.g., *visita medica* "medical examination"). Thus, we selected 84 collocations in total. Experimental items belonged to one of the three frequency bands; for each band we report the mean, standard deviation (*SD*), and range of MI values, as MI was used as a criterion in the identification of items (Table 2):

1. High frequency (14 V + N collocations, e.g., *apportare una modifica* "to make a change"; 14 N + Adj collocations, e.g., *lingua straniera* "foreign language");

2. Medium frequency (14 V + N collocations, e.g., *recepire un messaggio* “get the message”; 14 N + Adj collocations, e.g., *verità amara* “bitter truth”);
3. Low frequency (14 V + N collocations, e.g., *rispettare un orario* “to keep to a schedule”; 14 N + Adj collocations, e.g., *mente aperta* “open mind”).

To check whether the three frequency bands differed significantly, we ran an analysis of variance (ANOVA) with Tukey comparison. The results showed that the three frequency bands differed significantly from each other (ps $< .001$).

Measure of Directionality

Although it was not part of the item identification procedure, after the data had been collected and following the suggestion of one of the reviewers, we derived delta P scores (ΔP; Gries, 2013; Schneider, 2020) for each collocation and included them in the final analysis. Delta P measures how much one word attracts the other word within a collocation, taking into account directionality, as the attraction between the two words may not be symmetrical (Gries, 2013). With directionality taken into account, delta P produces two scores: $\Delta P_{forward}$ measures the extent to which “$word_1$ is more predictive of $word_2$ than vice versa” (Gries, 2013, p. 148), whereas $\Delta P_{backward,}$ measures the opposite—the extent to which $word_2$ more strongly predicts $word_1$ than the other way around. The two delta P scores[2] were calculated following Gries’s formula (2013, p. 144):

$$\Delta P_{forward} = p\,(word_2/word_1 = present) - p\,(word_2/word_1 = absent) \quad (1)$$

$$\Delta P_{backward} = p\,(word_1/word_2 = present) - p\,(word_1/word_2 = absent) \quad (2)$$

$\Delta P_{forward}$ is calculated as the difference between the probability (p) of $word_1$ occurring when $word_2$ is present and the probability of $word_1$ occurring when $word_2$ is absent. Conversely, $\Delta P_{backward}$ is calculated as the difference between the probability of $word_2$ occurring when $word_1$ is present and the probability of $word_2$ occurring when $word_1$ is absent.

Association Strength and Cloze Probability Norming Tasks

We conducted two word-association tasks in order to calculate the association strength of experimental items. Here, association strength refers to how much two collocates are likely to be entrenched in the speaker’s mental lexicon. In the first task, 30 L1 speakers of Italian not involved in the main experiment read the 42 verbs from the V + N collocations and then provided the three nouns that came first to mind (e.g., *mantenere* "to keep" > 1. _______; 2. _______; 3. _______) . In the second task, 30 L1 speakers of Italian not involved in the main experiment nor in the first word-association norming task, read the 42 nouns of the N + Adj collocations and then provided the three adjectives that came first to mind (e.g., *visita* "examination" > 1. _______; 2. _______; 3. _______). Association strength was calculated by giving 3 points to the first association response, 2 points to the second association response, and 1 point to the third one. Following this procedure, each collocation was given a mean association strength score ranging from 0 (the second word of the collocation, i.e., the noun in a V + N collocation or the adjective in a N + Adj collocation, was not provided by anyone) to 3 (the second word was provided by all participants). The average score for V + N collocations was 0.60 (range: 0–2.7); the average for N + Adj collocations was 0.72 (range: 0–2.7).

Following the above procedure, we conducted a cloze probability norming task in order to estimate the predictability of each target (i.e., how likely a reader is to complete each sentence with the target). Collocations were embedded in sentences, with the second term of the collocation removed (e.g., *Maria mantiene una* _______ "Maria keeps a ________"). In this task, 20 L1 speakers of Italian not otherwise involved in the experiment completed the sentence with the first word that came to mind. Each collocation could be given a predictability score from 0 (target not provided by anyone) to 1 (target provided by everyone). V + N collocations were completed with the noun target on average by 0.3 out of 20 participants (range: 0–1). Similarly, N + Adj collocations were completed with the adjective target on average by 0.3 (range: 0–1) out of 20 participants.

Lexical Decision Task

In the LDT, the verb (e.g., *mantenere* "to keep") of the V + N collocation (e.g., *mantenere* + *promessa* "to keep a promise") and the noun (e.g., *visita* "examination") of the N + Adj collocation (e.g., *visita medica* "medical examination") were used as primes. In contrast, the noun of the V + N collocation (e.g., *promessa* "promise") and the adjective of the N + Adj collocation (e.g., *medica* "medical") were used as targets.

The 84 V + N and N + Adj collocations were matched with 84 V + N and N + Adj control pairs. For each collocation, there was a control pair with the same prime but a different target. The control target was similar in length (± 1 character) and frequency to the collocation target (e.g., collocation *mantenere* + *promessa* "to keep a promise" vs. control *mantenere* + *armonia* "to keep harmony"). Although control pairs were attested in the reference corpus, they

Table 3 Frequency and mutual information bands of control pairs

	Frequency			Mutual information		
Band	*M*	*SD*	Range	*M*	*SD*	Range
High	7.6	4.6	1–10	0.5	0.2	0.1–0.8
Medium	8.8	4.4	1–11	0.3	0.4	0.1–0.8
Low	4.4	3.5	2–10	0.3	0.2	0.1–0.6

were infrequent and had a MI score of less than 3 (Table 3). Finally, the control targets were never produced as an associate of the prime in the word association task (see above).

Collocations and control pairs were combined into two counterbalanced lists. Each list contained 21 V + N collocations and 21 N + Adj collocations (seven V + N collocations and seven N + Adj collocations from each frequency band), together with 21 V + N control pairs and 21 N + Adj control pairs. Each prime and target was used only once in each counterbalanced list. A single set of 84 prime–nonword pairs (e.g., *mantenere* + *scovassa*) was created by using the pseudoword generator Wuggy (Keuleers & Brysbaert, 2010) and added to both lists. In total, each list contained 168 items (84 experimental word pairs and 84 nonword pairs).

Procedure

A LDT (Neely, 1991) was created and run in DMDX (Forster & Forster, 2003). Participants were randomly assigned to one of the two counterbalanced lists. They were tested individually in a psycholinguistic laboratory. Items were displayed on a computer screen, and responses were registered by a Logitech F310 Gamepad. Figure 1 shows the presentation sequence for the trials. Each trial started with a fixation point ("+") in the middle of the screen, which was displayed on the screen for 500 ms. The fixation point was replaced by a pattern mask (presented for 500 ms) as a set of hashtags (#######); the pattern mask was replaced by the prime, which stayed on the screen for 150 ms. Although the psycholinguistic literature indicates that prime word duration should be between 50 and 60 ms to avoid intervention of strategic processes, Jiang (2012) advises using longer prime durations when participants are L2 speakers. Thus, we presented primes for 150 ms. To ensure that participants would not be able to identify the prime, we extended the duration of the pattern mask to 500 ms (Jiang, 2012).

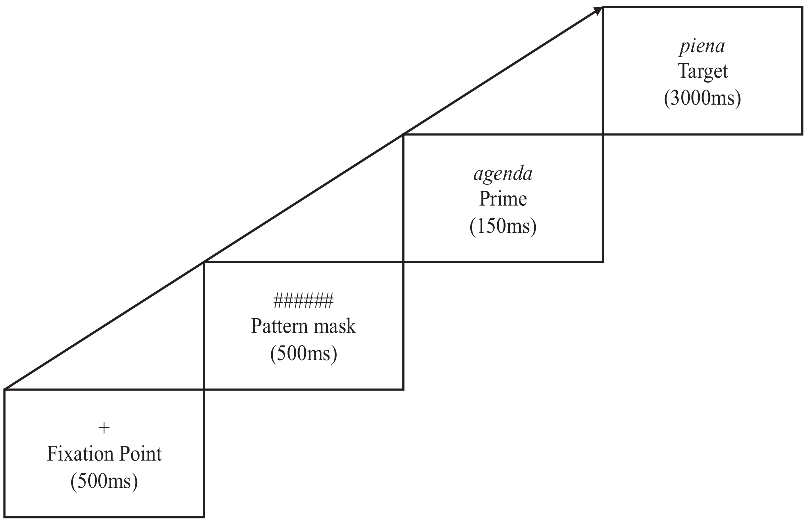

Figure 1 Presentation sequence in the primed lexical decision task.

The prime was immediately replaced by the target, and participants had 3,000 ms to respond. Following participants' response, the screen went blank for 1,000 ms before the onset of the next trial. A pattern mask was used to ensure the elicitation of automatic priming and to avoid intervention of strategic processes by participants (i.e., anticipation of the target; McNamara, 2005). Participants were instructed to press "YES" if they believed the string of letters was a real word in Italian, and to press "NO" if they believed the string was not a real word in Italian. They were asked to respond as quickly and as accurately as possible. RTs were measured from the onset of the target to participants' response. The task began with 32 practice trials. Items were presented in two blocks with a self-paced break between the two blocks. The order of the items was randomized.

Analysis and Results

We hypothesized that collocational priming should be found for both L1 and L2 speakers. We expected both groups of participants to respond faster to higher frequency collocations than to control conditions. Further, we did not expect to find evidence of priming for medium- and low-frequency items. Moreover, we expected L2 speakers' exposure to Italian to significantly affect their RTs.

Table 4 First language speakers' median reaction times (RTs) and error rates (ERs) for collocation and control

	RT (ms)			ER (%)	
Condition	Collocation	Control	Difference	Collocation	Control
High frequency	635	652	17	2.1	1.1
Medium frequency	642	621	−21	1.1	1.1
Low frequency	637	644	7	1.8	1.3

Table 5 Second language speakers' median reaction times (RTs) and error rates (ERs) for collocation and control

	RT (ms)			ER (%)	
Condition	Collocation	Control	Difference	Collocation	Control
High frequency	750	792	42	4.9	9.8
Medium frequency	766	796	30	7.1	5.8
Low frequency	769	751	−18	5.1	8.4

RTs faster than 200 ms, RTs slower than 2,000 ms, and items with a standard deviation of ±2 above and below the participant's RT mean were considered as outliers and removed prior to data analysis. This resulted in the loss of 5.3% of L1 data (collocation: 1.5%; control: 3.8%) and the loss of 5.8% of L2 data (collocation: 2.2%; control: 3.6%). Erroneous responses, including missing responses to real words, were excluded from the analyses (L1 speakers = 1.3 %, L2 speakers = 7%). Accuracy (calculated as the percentage of correct answers to real words) was 98.6% and 93% for L1 and L2 speakers, respectively. Tables 4 and 5 show RTs and error rates for L1 and L2 speakers.

We performed the analyses using mixed-effects modeling in R (e.g., Baayen et al., 2008). All continuous predictors were scaled to reduce collinearity. We built two models to investigate L1 and L2 speakers' RTs to the target, using R (Version 3.6.3, 2020-02-29) and the packages lme4 (Bates et al., 2012) and lmerTest (Version 2.0-6). We used ggplot2 (Wickham, 2016) to build model plots. The following predictors were included in the first model, which used L1 speakers' (logarithmically transformed) RTs as the dependent variable: (a) trial number (continuous); (b) condition (two-level categorical: collocation vs. control; reference level: control); (c) type (two-level categorical: V + N collocations vs. N + Adj collocations; reference level: N + Adj);

(d) log-adjusted phrase frequency (continuous); (e) frequency band (three-level categorical: high, medium, low; reference level: high); (f) MI score (henceforth, MI; continuous); (g) association strength (based on L1 speakers' responses collected in the two word association tasks; continuous); (h) predictability (based on L1 speakers' responses collected in the cloze probability task; continuous); (i) $\Delta P_{forward}$ (continuous); (j) $\Delta P_{backward}$ (continuous); and (k) length of the target item (in letters; continuous). Furthermore, the following interactions were included in the model: (a) Condition × MI; (b) Condition × Type; (c) Condition × Frequency Band; (d) Condition × Log-Adjusted Phrase Frequency; (e) Condition × Association Strength; and (f) Condition × Predictability.

In the second model, we used L2 speakers' (logarithmically transformed) RTs as the dependent variable. We used the same predictors as in the first model, with the addition of (a) proficiency (treated as the mean of the four skills: speaking, writing, listening, and reading) and (b) exposure (amount of time learners spent studying Italian in Italy, in months). Further, the following interactions were added to the ones used in the first model: (a) Condition × Proficiency; (b) Condition × Exposure; (c) Proficiency × Frequency Band; and (d) Proficiency × Log-Adjusted Phrase Frequency. Participants and items were included as random effects.

In both models we used a step-by-step backward model selection procedure (e.g., Manning, 2007), based on the Akaike information criterion (AIC) values, to select the most plausible model (Symonds & Moussalli, 2011). We started with the most complex model that included all the predictors and interactions of interest with a maximally specified random structure including by-participant and by-item intercepts and slopes as random effects (Barr et al., 2013), but due to convergence issues, models were simplified to by-participant and by-subject intercepts. We started removing one interaction term at a time by selecting the model with the lowest AIC value. We removed interactions and predictors only if, when removing them, the AIC value decreased, and if the decreases in AIC values were significant. At each step of this model selection process, we used likelihood ratio tests to compare pairs of models and to find the best fit (Baayen et al., 2008). We repeated the backward selection step procedure until no further interaction terms and predictors could be removed.

We checked the assumptions of the final models (linearity, normality of residuals, normality of random effects, and homogeneity of variance) by producing validation graphs. Finally, variance inflation factors (VIFs) were used to check multicollinearity: All VIF scores were smaller than 5. Comparisons between categorical variable levels (e.g., frequency bands) and interactions

Table 6 Summary of the model with first language speakers' reaction times

Fixed effects	*b*	*SE*	95% CI	*t*	*p*
Intercept	2.84	4.48	[−0.79, 0.06]	63.38	< .001
Condition (collocation)	1.29	5.46	[2.77, 2.83]	−2.36	.01
Length	−2.89	1.44	[−0.81, 0.01]	2.01	.05
Low-frequency band	0.38	1.91	[−0.70, 0.07]	−1.01	.36
Medium-frequency band	0.19	1.11	[−2.04, 0.04]	−1.72	.09
Phrase frequency	−1.93	1.32	[−0.68, 0.03]	−1.46	.14
MI	−9.78	1.74	[−2.89, 0.02]	1.87	.06
Predictability	−8.85	9.37	[−1.43, −0.03]	0.02	.03
$\Delta P_{forward}$	−1.52	−2.43	[−0.005, 0.004]	1.55	.95
$\Delta P_{backward}$	1.15	2.28	[−0.003, 0.005]	0.50	.61
Condition × Phrase Frequency	−4.66	1.88	[−4.49, 0.05]	2.47	.01
Condition × Predictability	−2.99	1.27	[−1.44, 0.07]	2.33	.02
Condition × Frequency Band	3.87	1.43	[−6.68, −0.01]	2.69	< .001

Random effects	Variance	*SD*	$R^2_{conditional}$	$R^2_{marginal}$
Item	0.17	0.01	.51	.08
Subject	0.53	0.07		

Note. Intercept levels: Condition (control). MI = mutual information; ΔP = delta P.

with other variables and their effect sizes were performed using the emmeans package (Version 1.4.2, 2019). Main effects and interactions were plotted using the ggplot2 package (Version 3.2.1, 2019).

L1 Speakers

Table 6 shows the final model for L1 speakers, the main effects, and their *p* values. The model explained 51% of variance; of this variance, fixed effects explained 8%. The main effect of condition (collocation) was found significant ($b = 1.29$, $t = -2.36$, $p = .01$), as was predictability ($b = -8.85$, $t = 0.02$, $p = .03$). Further, the following interactions were significant: Condition × Phrase Frequency ($b = -4.66$, $t = 2.47$, $p = .01$), Condition × Predictability ($b = -2.99$, $t = 2.33$, $p = .02$), and Condition × Frequency Band ($b = 3.87$, $t = 2.69$, $p < .001$).

Post hoc analysis revealed a priming effect for collocation compared to the control condition ($t = 0.28$, $p = .01$, $d = 0.25$, 95% CI [0.06, 0.08]). The significant interaction between condition and frequency band showed that L1 speakers' RTs were affected by frequency band. Post hoc analysis revealed a priming

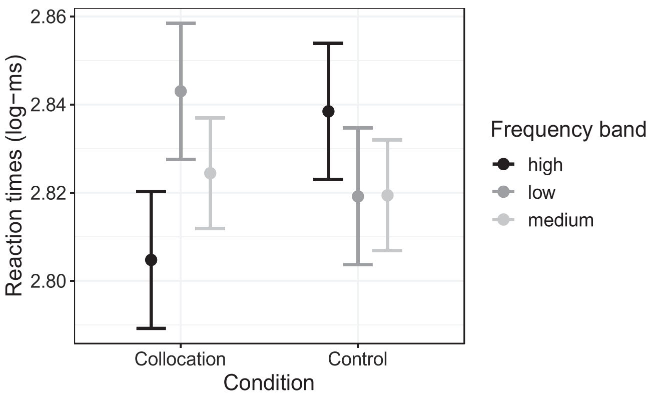

Figure 2 First language speakers' reaction times as a function of frequency band and condition. Error bars represent confidence intervals.

effect between the elements of high-frequency collocations compared to the elements of control pairs ($t = 1.32$, $p = .01$, $d = 0.33$, 95% CI [2.80, 2.85]). No priming effect was found between the elements of medium-frequency collocations relative to the control condition ($t = 0.01$, $p = .43$, $d = 0.10$, 95% CI [2.79, 2.84]), or between the elements of low-frequency collocations relative to their controls ($t = 0.04$, $p = .96$, $d = 0.03$, 95% CI [2.78, 2.82]). The interaction between frequency band and condition is plotted in Figure 2.

Finally, predictability interacted with condition and significantly affected L1 speakers' RTs (Figure 3). Further comparisons showed that as predictability increased, L1 speakers took less time to make a lexical decision ($t = 2.38$, $p = .01$, $d = 0.17$, 95% CI [0.20, 0.30]). Finally, the interaction between condition and phrase frequency was significant ($t = 2.27$, $p = .02$, $d = 0.11$, 95% CI [0.42, 0.62]): The strength of priming effect increased with the increase of phrase frequency (Figure 4).

L2 Speakers

Table 7 contains the final model for L2 speakers, the main effects, and their p values. The model explained 35% of variance; of this variance, fixed effects explained 7%. The following predictors were found to significantly affect L2 learners' RTs: condition (control) ($b = 2.91$, $t = 40.11$, $p < .001$), type ($b = 1.62$, $t = 2.30$, $p = .02$), phrase frequency ($b = -2.10$; $t = 2.07$, $p = .03$), length ($b = -6.07$, $t = 2.05$, $p = .04$), MI ($b = -1.19$, $t = 3.54$, $p < .001$), and exposure ($b = -8.76$, $t = 2.57$, $p = .01$). The interaction between condition

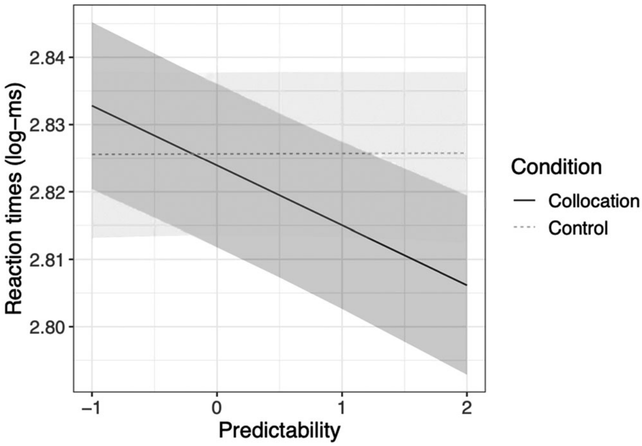

Figure 3 The effect of predictability on first language speakers' reaction times. Shared areas represent confidence intervals.

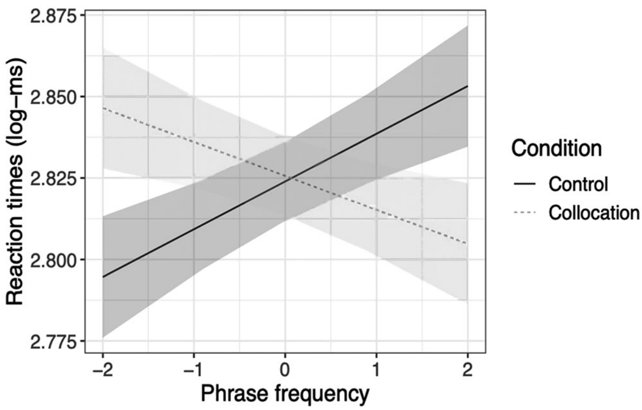

Figure 4 First language speakers' reaction times as a function of phrase frequency and condition. Shared areas represent confidence intervals.

and exposure was also significant ($b = -3.09$, $t = -2.29$, $p = .02$). Post hoc analyses revealed that the difference in L2 speakers' RTs between collocations and their control conditions was significant ($t = 1.01$, $p = .01$, $d = 0.14$, 95% CI [0.03, 0.05]). Moreover, phrase frequency had a negative effect on L2 learners' RTs (see Table 7): As phrase frequency increased, learners' RTs decreased. Indeed, a priming effect was found between high-frequency phrase constituents but not between low-frequency ones.

Table 7 Summary of the model with second language speakers' reaction times

Fixed effects	*b*	*SE*	95% CI	*t*	*p*
Intercept	2.91	7.25	[2.71, 3.01]	40.11	< .001
Condition (collocation)	−4.43	2.34	[−0.01, 9.02]	1.89	.06
Type	1.62	7.04	[0.02, 3.03]	2.30	.02
MI	−1.19	3.37	[0.01, 1.84]	3.54	< .001
Phrase frequency	−2.10	1.01	[0.02, 3.02]	2.07	.03
Length	−6.07	2.95	[0.01, 1.18]	2.05	.04
Proficiency	−2.39	1.90	[−0.06, 1.33]	1.25	.21
Association score	5.76	7.83	[−0.01, 1.51]	0.07	.94
Exposure	−8.76	3.40	[0.01, 1.54]	2.57	.01
$\Delta P_{forward}$	−6.04	5.15	[−0.01, −4.05]	−1.17	.24
$\Delta P_{backward}$	1.16	4.55	[−0.01, 1.01]	0.25	.80
Condition (Collocation) × Exposure	−3.09	1.34	[0.01, 1.82]	−2.29	.02
Condition (Collocation) × MI	−5.44	3.78	[−0.01, 2.01]	−1.43	.15
Condition (Collocation) × Association Score	1.78	9.89	[−0.03, 1.23]	−1.80	.07

Random effects	Variance	*SD*	$R^2_{conditional}$	$R^2_{marginal}$
Item	0.12	0.03	.35	.07
Subject	0.29	0.05		

Note. Intercept levels: Condition (control). MI = mutual information; ΔP = delta P.

Moreover, length significantly affected L2 learners' RTs, with shorter target items having overall shorter RTs. MI was a significant predictor—the priming effect increased as the MI values increased. In addition, unlike the RTs of L1 speakers, those of L2 speakers were significantly affected by the type of collocation. Post hoc analysis revealed that N + Adj items were processed faster than V + N items ($t = -3.23$, $p < .001$, $d = 0.13$, 95% CI [0.15, 0.19]) relative to their controls. The effect of type is plotted in Figure 5. Furthermore, exposure to L2 Italian significantly affected L2 learners' RTs (see Table 7): Learners' RTs became faster as their exposure to L2 increased. Moreover, the effect of L2 exposure significantly interacted with condition. Unexpectedly, L2 speakers with less exposure reacted faster to all experimental conditions (i.e., collocation and control) compared to learners with more exposure to L2. However, further analyses revealed that L2 learners with less exposure to Italian

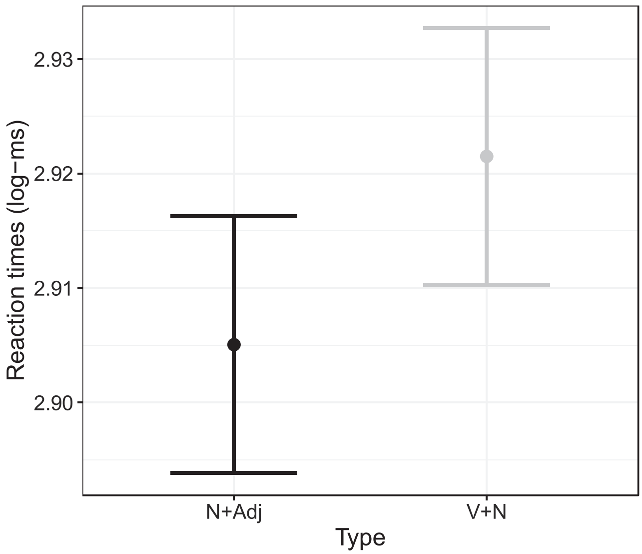

Figure 5 The effect of type of collocation on second language speakers' reaction times. Error bars represent confidence intervals.

responded to collocations and their respective controls in a comparable way. In contrast, L2 speakers with more exposure to Italian responded faster to collocations than to controls ($t = 2.31$, $p = .02$, $d = 0.18$, 95% CI [2.13, 2.19]). The interaction between condition and exposure to L2 is shown in Figure 6.

General Discussion

In the present study, we sought to provide further insights into the phenomenon of collocational priming in L1 and L2 speakers of Italian. In what follows below, we discuss our posed research questions one at a time.

Research Question 1: Does Collocational Priming Occur in L1 and L2 Speakers of Italian?

We found evidence of collocational priming in both L1 and L2 speakers of Italian. Both groups of participants responded faster to collocation targets (*mantenere* + *promessa* "to keep + promise") than to control targets (*mantenere* + *armonia* "to keep + harmony"). This is in line with the results previously reported in the literature (e.g., Cangır et al., 2017; Durrant & Doherty, 2010; Ellis et al., 2008, 2009). This finding provides support to Hoey's (2005) claim that collocations are "a psychological association between words" (p. 5). When

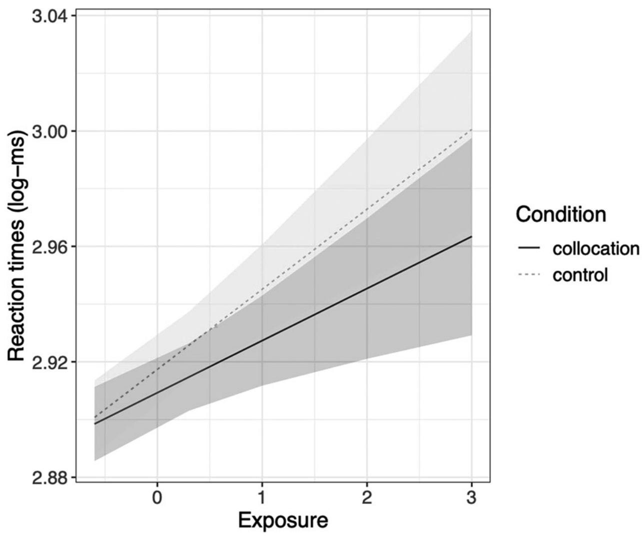

Figure 6 Second language speakers' reaction times as a function of amount of exposure to the second language (in months) and condition. Shared areas represent confidence intervals.

our participants encountered the first element of a collocation (e.g., *mantenere* "to keep"), they expected, in cognitive terms, the second one (e.g., *promessa* "promise"). When their expectation was met, they responded faster than when it was not met (e.g., *armonia* "harmony"). Our finding offers support to Hoey's (2005) theory of a relationship between collocations in text and collocations in mind, since our collocations were extracted from a large corpus of Italian using a frequency-based approach. Thus, we can argue that corpus-based measures are a valid method for identifying collocations that have psycholinguistic reality (see also Durrant & Doherty, 2010). Our results imply that highly frequent and strongly associated (e.g., in terms of MI) collocations extracted from a corpus are likely to be entrenched in the mental lexicon. Therefore, identifying collocations in texts using corpus-derived association measures (e.g., MI) allows researchers to extract reliable stimuli for investigating how collocations are represented in both L1 and L2 mental lexicons, given that high frequency and strongly associated collocations "are likely to be good targets for language learning" (Durrant & Doherty, 2010, p. 145).

It is important to note that we found no effect of association strength. This is different from the findings of Durrant and Doherty (2010, Experiment 2), who observed a significant effect of priming between collocations that were also normative associates. In contrast, collocational priming in our study did not depend on whether or not collocation constituents were associates in terms of free association norms. However, it is difficult to separate collocations from associates—there are few normative associates that are not also frequent collocations. It is likely that frequent collocations that elicit priming are also normative associates.

Research Question 2: Does Phrase Frequency Modulate the Strength of Collocational Priming?

The second aim of the present study was to investigate the tenet that frequency modulates the strength of collocational priming (e.g., Cangır et al., 2017; Durrant & Doherty, 2010). Indeed, both L1 and L2 speakers' RTs were significantly affected by frequency. On the one hand, we found that the interaction between frequency bands and condition (collocation vs. control) significantly affected L1 speakers' RTs. A priming effect was found between high-frequency collocations and their control conditions. In contrast, no priming effect was found between medium- and low-frequency collocations and their respective control pairs. On the other hand, L2 speakers' RTs were significantly affected by phrase frequency. Learners' RTs became faster as the frequency of collocation increased. This finding is in line with those of previous studies (e.g., Cangır et al., 2017; Durrant & Doherty, 2010; Ellis et al., 2008, 2009), confirming that high-frequency co-occurrences are more likely to be psychologically salient than low-frequency co-occurrences.

Although the issue is not part of Research Question 2, it is worth noting that MI was not a significant predictor of L1 RTs, suggesting that L1 collocational priming may be modulated by phrase frequency alone. In contrast, L2 speakers' responses were significantly affected by MI scores. This finding is in line with Öksüz et al. (2020), who found that the processing speed of L2 speakers was similarly affected by association measures, but it goes contrary to the results of Ellis et al. (2008), who found that only L1 speakers were affected by MI. We should be careful about attributing an important role to MI in modulating the mechanisms of L2 collocational priming. As Cangır et al. (2017) note, MI did not reflect a strong association between collocation elements, as it is highly sensitive to low-frequency word pairs. Further, Öksüz et al. (2020) also compared different association measures and their role in collocation processing. Their findings showed that the model including logDice[3] scores was a better fitting model than the model with MI scores. This might be due to the fact that, unlike MI, logDice does not overestimate low-frequency co-occurrences (e.g., Gablasova et al., 2017; Öksüz et al., 2020). For example, *social policy* and *annual report* have frequencies of 876 and 641, respectively; however, the former has a MI score of 3.74 whereas the latter has a MI score of 5.78, with both collocations having similar logDice values of 7.19 and 7.13, respectively (frequencies extracted from the *British National Corpus*, BNC XML edition).

A few words need to be said about the effect of the delta P measure. The results showed a trend among both L1 and L2 participants to respond more quickly as $\Delta P_{forward}$ increased (see Tables 6 and 7), in contrast to $\Delta P_{backward}$. This aligns with the directional nature of our priming. However, $\Delta P_{forward}$ did not significantly influence RTs, contrary to findings in previous studies (e.g., Cangır et al., 2017; Cangır & Durrant, 2021). The divergence from prior research lies in the respective experimental designs; our study employed a constant prime with a changing target, whereas previous studies used a constant target with a changing prime. In experimental designs similar to ours, where various variables might influence the effect of priming (e.g., lexical properties of the target; Jiang, 2012), delta P may not significantly modulate priming effects. Conversely, in other studies (e.g., Cangır et al., 2017; Cangır & Durrant, 2021), where the difference in RTs is more likely to depend on the relationship between the prime and the target, the role of delta P may be more apparent.

Research Question 3: Do Exposure to the L2 and L2 Proficiency Play a Role in Strengthening the Associative Links Between Collocation Constituents?

Proficiency did not significantly modulate L2 speakers' RTs, whereas exposure to the L2 did. L2 speakers with a greater amount of exposure to Italian responded faster to collocation targets than to control targets, compared to L2 speakers with less exposure. Interestingly, L2 learners with a greater amount of exposure to the L2 responded more *slowly* to both collocations and control items compared to learners with less exposure to Italian. It is noteworthy that during the design stage, two options were available to us: (a) keeping the prime constant and using different targets; and (b) keeping the target constant and using different primes (see Jiang, 2012, for a discussion of the two possibilities). We adopted the first approach, as we were interested in investigating whether the first element of a collocation activates the second one, as per Hoey's (2005) theory. However, it is possible that participants' responses were modulated not only by the relationship between the prime and the target, but also by such lexical properties as the number of different collocates available for the first element presented (Jiang, 2012). It is possible that reading the verb in a V +

N collocation led to different collocates becoming activated, not only the target noun. The learners with less exposure to Italian in our study were likely to have had a smaller vocabulary size and thus a smaller set of potential collocates for each verb than the learners with more exposure. With a smaller set of possible collocates that could be activated, learners with less exposure to the L2 responded faster to targets compared to learners with more exposure, who were likely to have had a larger set of collocates for each verb that could be searched and become activated as they read the first element of a collocation. We thus tentatively take our results to suggest that having a larger set of collocates may result in more time being required to process the second element of a collocation when compared to having a smaller set of collocates available. This tenet, however, requires further investigation.

Although L2 learners with less exposure to Italian read the target items faster than their counterparts with more exposure, they nevertheless responded to collocations and controls in a similar way. In contrast, learners with more exposure were reliably faster to respond to collocations than to their controls. Thus, exposure to the L2 does affect the processing of collocations: L2 speakers with more exposure may in fact process collocations in a similar way to L1 speakers. This finding offers support to Hoey's (2005) suggestion that there may be little difference in the mechanisms responsible for L1 versus L2 collocational priming—the key lies in the amount of exposure to language. If L2 speakers are sufficiently exposed to the L2, the priming mechanisms at the root of the acquisition of L2 collocations may be comparable to those at the root of the acquisition of L1 collocations.

We now turn to the effect of L2 proficiency, which was found nonsignificant. In our study, proficiency was measured as the mean of self-reported speaking, writing, listening, and reading. Although this is a common way to assess proficiency (e.g., Siyanova-Chanturia et al., 2011; Sonbul & Schmitt, 2013), some researchers caution that this approach may lead to proficiency being under- or overestimated (Ma & Winke, 2019). Importantly, the collocational priming theory as conceived by Hoey (2005) is based on the idea of the entrenchment of connections between collocation elements, with entrenchment being due to exposure to the L2 rather than contingent on L2 proficiency. Clearly, exposure to the L2 (operationalized as time spent in L2 country) and L2 proficiency are distinct measures that should be treated and analyzed separately. Our results suggest that the two measures may differently predict collocational knowledge and collocation processing. This is particularly important because the processing of collocations may be more dependent on the exposure to a naturalistic L2 environment than on the time spent in a L2 classroom (e.g., some higher proficiency L2 learners in our pool had spent no time in Italy at all). However, more research is needed to tease these two apart, with proficiency operationalized both through self-reported measures and in a more objective way (e.g., established language certification).

Research Question 4: To What Extent Are the Mechanisms at the Root of L2 Collocational Priming Comparable to Those Responsible for L1 Collocational Priming?

Finally, we found that L1 speakers and L2 speakers with greater exposure to the L2 processed collocations in a comparable way in terms of RTs. However, two differences emerged between L1 and L2 collocational priming, concerning the effect of predictability and the effect of the type of collocation (V + N vs. N + Adj). We found a significant interaction between condition (collocation vs. control) and predictability, but only for L1 speakers, suggesting that phrase frequency is not the only variable determining the representation of collocations in the mental lexicon. This is not surprising, as predictability has long been argued to be an intrinsic property of multiword expressions (Siyanova-Chanturia & Omidian, 2020; Siyanova-Chanturia et al., 2017). Interestingly, predictability did not affect L2 speakers' RTs. In our study, norming scores for predictability were collected from L1 speakers (as is customary in phrasal processing research). However, it may be that instances that are highly predictable for L1 speakers may not be predictable (or may be less predictable) for L2 learners. Future researchers should consider using predictability scores collected from both L1 and L2 speakers.

Further, the type of collocation (V + N vs. N + Adj) significantly affected the RTs of L2, but not L1, speakers. L2 speakers responded faster to N + Adj than to V + N collocations, relative to their controls. This result is contrary to that of Cangır et al. (2017), whose L1 participants responded faster to V + N collocations than to Adj + N collocations, relative to their controls. Cangır et al.'s (2017) results showed that both categories elicited a significant priming effect compared to control conditions. However, the amount of priming was greater for V + N collocations than for Adj + N collocations, suggesting that nouns may be processed faster when they are primed by a verb than when they are primed by an adjective (Cangır et al., 2017). A possible explanation for the discrepancy between the two studies may lie in the degree of concreteness of primes. Turkish and Italian represent two different typologies. In the former, the typical word order is Adj + N, whereas in the latter it is N + Adj. Therefore, in Cangır et al.'s (2017) study, the adjective was found to prime the noun, unlike what was found in our study. It is well known that concreteness affects RTs in priming experiments (Jiang, 2012); adjectives are less concrete than nouns and this could lead to the inhibition of priming, as targets (i.e., nouns) may be less readily recalled. In contrast, nouns, which are more concrete compared to adjectives, may facilitate the activation of the target words (e.g., adjectives).

Our finding is in line with that of Cangır and Durrant (2021), who observed a greater amount of crosslinguistic priming for Adj + N collocations compared to V + N collocations. In our study, L2 learners responded faster to N + Adj collocations than to V + N collocations, suggesting that the two phrase types may be represented differently in the L2 mental lexicon. This might be due to the fact that Italian V + N collocations present more variability than Italian N + Adj collocations. For example, the verb in V + N collocations can occur in many different forms (e.g., *Lei mantiene/ha mantenuto/manteneva/manterrà la promessa* "She keeps/has kept/kept/will keep a promise"). Further, V + N collocations may vary in the use of the determiner (e.g., *commettere un errore* "to make a mistake" vs. *perdere tempo* "to waste time"), and other words can be inserted between the two constituents (e.g., *ho perso tempo* "I have wasted time" → *ho perso troppo tempo* "I have wasted too much time"). Because V + N collocations present a large amount of variability and are conjugated in Italian, L2 learners are unlikely to encounter V + N collocations always in the same form and may thus perceive the two collocation elements as less strongly associated than the two elements in N + Adj collocations, which are less affected by variability. This tenet may explain why L2 learners of Italian responded to the targets of N + Adj collocations faster than to the targets of V + N collocations.

Another possible explanation for the difference between L1 and L2 speakers in the processing of the two types of collocations (V + N and N + Adj) may lie in the variation in word order between the learners' (various) L1s and Italian. Because our participants came from a variety of L1s, it is plausible that some L1s may exhibit word order patterns different from those of Italian. For instance, if the noun precedes the verb in the learner's L1, this could account for why learners took longer to respond to V + N collocations compared to N + Adj collocations. Indeed, the word order in collocations has been found to have a significant effect on collocational priming (Cangır & Durrant, 2021).

Future Directions and Limitations

Our results reveal that there is much to be gained from integrating corpus data with psycholinguistic approaches in the investigation of the cognitive reality of collocations. We found evidence of priming in collocations identified by a corpus-based approach (i.e., via frequency and MI). Our findings showed that only high-frequency collocations elicit the priming effect. Although the impact of frequency on the elicitation of collocational priming is well attested (Cangır et al., 2017; Cangır & Durrant, 2021; Durrant & Doherty, 2010), the role of association measures has been less investigated. In the present study, MI, a commonly used association strength measure, was used as a criterion for identifying collocations in the corpus. Our results showed that L1 speakers' RTs were not predicted by MI. In contrast, L2 speakers' responses were significantly affected by MI scores. However, other association measures (e.g., logDice, delta P) have been found to influence collocation processing (e.g., Öksüz et al., 2020). Future research could investigate how different association measures may modulate collocational priming. In addition, we should note that our collocations were extracted from a written corpus and priming was examined in a written mode. Thus, our findings are limited to the written modality, and future research should explore whether modality (written vs. spoken) may modulate L1 and L2 collocational priming.

Before we conclude, the following points ought to be noted. First, unlike earlier research (e.g., Cangır et al., 2017; Durrant & Doherty, 2010), in our study, primes were kept constant while targets were different. This was done because our goal was to test Hoey's (2005) theory of collocational priming. Therefore, our aim was to test whether, after reading the prime (the first element of the collocation), participants were likely to recall the target (the second element of the collocation). However, we cannot rule out an intervening effect of the lexical properties of the targets. Future researchers may consider comparing both experimental variations (i.e., keeping the prime constant while changing the target vs. keeping the target constant while changing the prime) to investigate in which of the two cases a stronger priming effect is elicited.

Second, we presented the prime for longer (i.e., 150 ms) than the duration typically used in the literature (i.e., 50 ms). This was done following Jiang's (2012) suggestion that caution needs to be exercised when presenting very short primes to L2 participants, because L2 words tend to take longer to process than L1 words. Thus, to minimize the likelihood of participants engaging in strategic processes during target recognition, we extended the duration of the pattern mask from 150 ms to 500 ms. Still, we cannot rule out the possibility of the intervention of strategic processes used by participants. Future research is needed to further validate the duration of prime in the automatic priming paradigm with L2 participants.

Third, as was noted above, participants' L1 backgrounds were heterogeneous. This prevented the consideration of other variables such as congruency and word order in both the L1 and the L2. Given that learners' L1 can have a significant effect on collocational priming (Cangır & Durrant, 2021), future research may consider testing learners from the same L1 background.

Finally, a potential application of the priming paradigm in L2 research is the examination of lexicalization in a L2. The priming paradigm allows researchers to investigate the automatic activation and processing of lexical information in a L2. By exposing learners to certain primes, researchers can observe how quickly and efficiently they retrieve and recognize related lexical items. Additionally, investigation of priming may help in uncovering the extent to which L2 learners have integrated new words or collocations into their mental lexicon. The faster and more accurately learners respond to primed words (or whole collocations), the stronger the evidence is for the successful lexicalization of these items in the learners' L2.

Conclusion

To conclude, the present study sought to provide new insights into the phenomenon of collocational priming. We carried out a LDT with the aim of eliciting automatic priming in L1 and L2 speakers of Italian. We take our findings to support Hoey's (2005) theory of collocational priming and the tenet that collocations have cognitive reality. Furthermore, we found evidence of collocational priming only between high-frequency collocation elements, which supports Hoey's (2005) claim that higher frequency collocating items are more likely to be mentally primed with each other than lower frequency ones. More generally, our results support Hoey's (2005) claim that the mechanisms at the root of L2 collocational priming are comparable to the mechanisms at the root of L1 collocational priming. Both L1 and L2 speakers responded faster to higher frequency collocating items than to their controls. Thus, the mechanisms underlying collocational priming may be similar for L1 and L2 speakers; the key is likely to be the amount of exposure to the language concerned. Indeed, the RTs of L2 speakers with more exposure to the L2 were comparable to those of L1 speakers, but differed from the RTs of L2 learners with less exposure to the L2.

Final revised version accepted 15 May 2024

Notes

1 All the materials can be accessed on the OSF website (https://osf.io/f7dm2/?view_only).

2 Although our priming is directional, we included both delta P scores in order to test the possible bidirectional activation of collocation links (see also Cangır et al., 2017).

3 logDice is a standardized measure operating on a scale with a fixed maximum value of 14 that expresses the tendency of two words to co-occur together relative to their frequency in the corpus (Gablasova et al., 2017).

References

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

Baroni, M., Bernardini, S., Comastri, F., Piccioni, L., Volpi, A., Aston, G., & Mazzoleni, M. (2004). Introducing the La Repubblica Corpus: A large, annotated, TEI (XML-)-compliant corpus of newspaper Italian. In *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC'04)*, Lisbon, Portugal: European Language Resources Association (ELRA).

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep in maximal. *Journal of Memory and Language*, *68*(3), 255–279. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D. M., Maechler, M., & Bolker, B. (2012). *lme4: Linear mixed-effects models using Eigen and S4* (R package; Version 1.0-6) [Computer software]. https://cran.r-project.org/web/packages/lme4/lme4.pdf

Cangır, H., Büyükkantarcıoğlu, S. N., & Durrant, P. (2017). Investigating collocational priming in Turkish. *Journal of Language and Linguistic Studies*, *13*(2), 465–486.

Cangır, H., & Durrant, P. (2021). Cross-linguistic collocational networks in the L1 Turkish–L2 English mental lexicon. *Lingua*, *258*, Article 103057. https://doi.org/10.1016/j.lingua.2021.103057

Du, L., Siyanova-Chanturia, A., & Elgort, I. (2023). Cross-language influences in the processing of L2 multi-word expressions. In I. Elgort, A. Siyanova-Chanturia, & M. Brysbaert (Eds.), *Cross-language influences in bilingual processing and second language acquisition* (pp. 187–210). John Benjamins.

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

Durrant, P., & Siyanova-Chanturia, A. (2015). Learner corpora and psycholinguistics. In S. Granger, G. Gilquin, & F. Meunier (Eds.), *The Cambridge handbook of learner corpus research* (pp. 57–77). Cambridge University Press.

Eguchi, M., & Kyle, K. (2023). Lexical characteristics of young L2 English learners' narrative writing at the start of formal instruction. *Journal of Second Language Writing*, *60*, Article 100975. https://doi.org/10.1016/j.jslw.2023.100975

Ellis, N. C., Frey, E., & Jalkanen, I. (2009). The psycholinguistic reality of collocation and semantic prosody (1): Lexical access. In U. Römer & R. Schulze (Eds.), *Exploring the lexis-grammar interface* (pp. 89–114). John Benjamins.

Ellis, N. C., Simpson-Vlach, R., & Maynard, C. (2008). Formulaic language in native and second-language speakers: Psycholinguistics, corpus linguistics, and TESOL. *TESOL Quarterly*, *41*(3), 375–396. https://doi.org/10.1002/j.1545-7249.2008.tb00137.x

Evert, S. (2008). Corpora and collocations. In A. Ludeling & M. Kyto (Eds.), *Corpus linguistics: An international handbook*. Mouton de Gruyter (pp. 1212–1248).

Firth, J. R. (1957). A synopsis of linguistic theory, 1930–55. In *Studies in linguistic analysis* (pp. 1–32). Oxford: Special volume of the Philological Society.

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, *35*(1), 116–124. https://doi.org/10.3758/BF03195503

Gablasova, D., Brezina, V., & McEnery, T. (2017). Collocations in corpus-based language learning research: Identifying, comparing, and interpreting the evidence. *Language Learning*, *67*, 155–179. https://doi.org/10.1111/lang.12225

Gries, S. T. (2013). 50-something years of work on collocations: What is or should be next … *International Journal of Corpus Linguistics*, *18*(1), 137–166. https://doi.org/10.1075/ijcl.18.1.09gri

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

Hoey, M. (1991). *Patterns of lexis in text*. Oxford University Press.

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

Jiang, N. (2012). *Conducting reaction time research in second language studies*. Routledge.

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods*, *42*, 627–633. https://doi.org/10.3758/BRM.42.3.627

Lenci, A., Lapesa, G., & Bonasinga, G. (2012). LexIt: A computational resource on Italian argument structure. In *Proceedings of the Eight International Conference on Language Resource and Evaluation (LREC'12)* (pp. 3712–3718). Istanbul, 23–25 May 2012.

Ma, W., & Winke, P. (2019). Self-assessment: How reliable is it in assessing oral proficiency over time? *Foreign Language Annals*, *52*(1), 66–86. https://doi.org/10.1111/flan.12379

Manning, C. (2007). *Generalized linear mixed models*. Unpublished course handout Standford University.

McNamara, T. P. (2005). *Semantic priming: Perspectives from memory and word recognition*. Psychology Press.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, *90*(2), 227–234. https://doi.org/10.1037/h0031564

Nation, I., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, *31*, 9–13. https://doi.org/10.26686/wgtn.12552197.v1

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In G. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264–336). Lawrence Erlbaum.

Öksüz, D., Brezina, V., & Rebuschat, P. (2020). Collocational processing in L1 and L2: The effects of word frequency, collocational frequency and association. *Language Learning*. Advance online publication. https://doi.org/10.1111/lang.12427

Schneider, U. (2020). ΔP as a measure of collocation strength: Considerations based on analyses of hesitation placement in spontaneous speech. *Corpus Linguistics and Linguistic Theory*, *16*(2), 249–274. https://doi.org/10.1515/cllt-2017-0036

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

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), 776–784. https://doi.org/10.1037/a0022531

Siyanova-Chanturia, A., Conklin, K., Caffarra, S., Kaan, E., & van Heuven, W. J. B. (2017). Representation and processing of multi-word expressions in the brain. *Brain and Language*, *175*, 111–122. https://doi.org/10.1016/j.bandl.2017.10.004

Siyanova-Chanturia, A., & Omidian, T. (2020). Key issues in researching multiword items. In S. Webb (Ed.), *The Routledge Handbook of Vocabulary Studies* (pp. 529–544). Routledge.

Siyanova-Chanturia, A., & van-Lancker Sidtis, D. (2019). What on-line processing tells us about formulaic language. In A. Siyanova-Chanturia, & A. Pellicer-Sánchez (Eds.), *Understanding formulaic language: A second language acquisition perspective* (pp. 38–61). Routledge.

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

Sonbul, S., & El-Dakhs, D. (2020). Timed versus untimed recognition of L2 collocations: Does estimated proficiency modulate congruency effects? *Applied Psycholinguistics*, *41*(5), 1197–1222. https://doi.org/10.1017/S014271642000051X

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

Symonds, M. R. E., & Moussalli, A. (2011). A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike's

information criterion. *Behavioral Ecology and Sociology*, *65*(1), 13–21. https://10.1007/s00265-010-1037-6
Toomer, M., & Elgort, I. (2019). The development of implicit and explicit knowledge of collocations: A conceptual replication and extension of Sonbul and Schmitt (2013). *Language Learning*, *69*(2), 1–35. https://10.1111/lang.12369
Wickham, H. (2016). *ggplot: Elegant graphics for data analysis*. Springer-Verlag.
Wolter, B., & Gyllstad, H. (2013). Frequency of input and L2 collocational processing. *Studies in Second Language Acquisition*, *35*(3), 451–482. https://10.1017/S0272263113000107

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Acessible Summary