Lexical fixedness and compositionality in L1 speakers' and L2 learners' intuitions about word combinations: Evidence from Italian

Irene Fioravanti (ID)
Università per Stranieri di Siena, Italy

Marco Silvio Giuseppe Senaldi
McGill University, Canada

Alessandro Lenci
Università di Pisa, Italy

Anna Siyanova-Chanturia
Te Herenga Waka – Victoria University of Wellington, New Zealand; Ocean University of China, China

Abstract
The present investigation focuses on first language (L1) and second language (L2) speakers' sensitivity to lexical fixedness and compositionality of Italian word combinations. Two studies explored language users' intuitions about three types of word combinations: free combinations, collocations, and idioms. In Study 1, Italian Verb+Noun combinations were embedded in sentential contexts, with control conditions created by substituting the verb with a synonym. L1 and L2 speakers rated sentence acceptability. In Study 2, the original verb was removed from sentences. Participants chose the verb from the list provided they felt was most acceptable. Computational measures were used to measure compositionality of word combinations. Mixed-effects modelling revealed that L1 and L2 speakers judged target word combinations differently in terms of lexical fixedness. In line with phraseological models, L1 speakers judged the use of a synonym as less acceptable in collocations than free combinations. On the contrary, L2 learners judged the use of a synonym as more acceptable in collocations than free combinations.

Corresponding author:
Irene Fioravanti, Università per Stranieri di Siena, Piazza Carlo Rosselli, 27, Siena, 53100, Italy
Email: irene.effe@gmail.com

However, all participants perceived idioms as least flexible of the three combination types. Results further showed an interesting effect of compositionality on the speakers' intuitions about the use of word combinations. Taken together, the findings provide new insights into how L1 and L2 speakers perceive word combinations that vary along the continua of lexical fixedness and compositionality.

Keywords

collocations, compositionality, free combinations, frequency, idioms, Italian, lexical fixedness, L2 learners

I Introduction

It is widely acknowledged that word combinations – collocations, idioms, binomials, and other word sequences – play an important role in the learning and use of a second language (Wray, 2002). For example, knowledge of such items has been shown to positively impact on second language (L2) learners' fluency (Howarth, 1998).

Because of the key role they play in L2 learning, word combinations have received a considerable amount of attention in L2 acquisition research in the past 20 years. Traditionally, research on word combinations has employed either a frequency-based or a phraseological approach. The former defines word combinations as frequent occurrences of two or more words in written or spoken discourse (Evert, 2008; Firth, 1957; Sinclair, 1991). The latter identifies phraseological units on the basis of compositionality and lexical fixedness (Cowie, 1981; Howarth, 1998). Despite conceptual and methodological differences, proponents of both approaches have found word combinations – collocations, in particular – to be problematic for L2 learners (Henriksen, 2013). In fact, collocations have been reported to be challenging even for advanced learners. As Laufer and Waldman (2011) note, although advanced learners produce more collocations compared to beginner and intermediate learners, they also err more frequently and tend to produce more deviant collocations compared to lower proficiency learners. As they posited, 'the development of collocation use is slow and uneven' (Laufer and Waldman, 2011: 664).

Previous studies have mostly dealt with collocation usage and the development of collocational knowledge (Durrant and Schmitt, 2009; Singleton et al., 2007; Siyanova and Schmitt, 2008; Sonbul, 2013). In addition, research has looked at the effect of frequency and congruency on the collocation processing (Sonbul, 2015; Wolter and Gyllstad, 2013; Wolter and Yamashita, 2018), and non-literal language processing and use in a L2 (Carrol and Conklin, 2017; Siyanova-Chanturia et al., 2011; Underwood et al., 2004). Most of these studies have focused on one type of word combinations at a time. In fact, few studies have to date investigated different types of phraseological units in one study using the same methodology. In the present investigation, we focus on first language (L1) and second language (L2) speakers' intuitions about three types of word combinations: free combinations, collocations and idioms. Our definition of these three types of phraseological units, based on semantic compositionality and lexical rigidity, draws on Ježek's (2005) and Masini's (2009) phraseological models.

The aim of the present study was to investigate L1 and L2 speakers' intuitions about the lexical fixedness of word combinations, which is one of the most commonly used criteria for classifying and distinguishing between them (Nesselhauf, 2003). Yet, to our knowledge, it has not been extensively investigated.

I Lexical fixedness and compositionality within the phraseological tradition

Phraseology is the study of the grammatical structure, meaning and use of word combinations (Cowie, 1998). The phraseological approach distinguishes word combinations in terms of lexical fixedness and compositionality. The former refers to whether or not a word within a phrase can be substituted with a synonym, while the latter has to do with how much the meanings of individual words contribute to the overall meaning of a phrase. Starting with Cowie (1994) and Howarth (1998), word combinations have often been placed along a continuum, with free combinations and idioms being at the far ends, and collocations somewhere in-between. Cowie (1994) and Howarth (1998) described word combinations on the basis of the criteria of substitution and semantic transparency. At one end of the continuum are free combinations (*blow a trumpet*) that are viewed as the least restricted, literal combinations, whose individual components can be freely combined and substituted. At the other end are idioms that are described as the most restricted word combinations. Idioms are often divided into figurative idioms, which have a literal and a figurative meaning (*blow your own trumpet*), and pure idioms, highly fixed and opaque word combinations that are only used figuratively (*blow the gaff*) (see Howarth, 1998). The third type of word combinations – collocations – are presented as lexically restricted word sequences (Cowie, 1998; Howarth, 1998) that allow a limited substitution of some of the components (*blow a fuse*). Despite their greater compositionality compared to idioms, collocation components can be combined only with a limited number of words (known as collocates).

Unlike Cowie (1981) and Howarth (1998), Hausmann (1989) and Mel'čuk (1998) focused on the relationship between the individual components of a combination, arguing that they do not enjoy the same status. With respect to collocations (*heavy smoker*), they claim, one of the elements is semantically independent (*smoker*), while the other (*heavy*) is 'selected' on the basis of the first element to express a specific meaning, that is, someone who smokes a lot (Mel'čuk, 1998).

Following this approach, Nesselhauf (2003) classified Verb+Noun collocations drawing on her criterion of 'restricted sense'. She considered the meaning of a verb (or noun) as restricted only if it is so specific that it can be combined with a very limited number of nouns (or verbs) and cannot be used in a given sense with all syntactically and semantically possible nouns (or verbs). According to these criteria, the components of a free combination are not used in a restricted sense, as they can be freely combined on the basis of their individual meanings (*read* and *book* in the expression *read a book*). Instead, the noun of a collocation is used in an unrestricted sense, but the sense of the verb is restricted (*take* in *take a picture*). The verb used in this specific sense can be combined with a limited number of nouns (*take a screenshot, take a selfie, take a shower*). The

components of an idiom are used in a restricted sense and their substitution is not possible (*smell* and *rat* in *smell a rat*).

Finally, Ježek (2005) and Masini (2009) proposed two additional classifications for Italian, which is our target language. Ježek (2005) based her classification on the criteria of semantic restriction, compositionality, lexical restriction and syntactic flexibility. She identified free combinations as compositional, lexically unrestricted and syntactically flexible. The constituents within a free combination can be freely replaced with synonyms and are hence not used in a restricted sense. Collocations (or restricted combinations) are defined as compositional combinations that are syntactically flexible but lexically restricted. Collocations do not allow free substitution of their individual components. They are subject to lexical restriction such that the choice of a specific word – the collocate – is dependent on another word – the base – to express a given meaning. Finally, idioms are defined as non-compositional combinations that are syntactically inflexible and lexically restricted.[1]

In Masini's (2009) classification, the distinction is based on the criteria of paradigmatic fixedness, syntagmatic fixedness and familiarity. Following these three criteria, Masini (2009) identified multiword expressions (i.e. idioms), collocations and preferential combinations. She defined multiword expressions (*fare attenzione* 'to pay attention') as frequent, familiar and both paradigmatically and syntagmatically fixed combinations. On the contrary, collocations (*prendere una decisione* 'make a decision') and preferential combinations (*gravemente malato* 'seriously ill') are the result of lexeme co-selection. The use of one lexeme can 'preferably' require the use of another lexeme, as in the case of preferential combinations (*caldo soffocante* 'sweltering heat'), or it can 'necessarily' require the use of another lexeme, as in the case of collocations (*aprire un conto* 'to open a bank account') (see Masini, 2009: 82). Moreover, collocations are characterized by paradigmatic fixedness, with the substitution of their components being restricted. In comparison, preferential combinations are more familiar and frequent than the other combinations, with the substitution of their components not being restricted.

2 Acquisition and processing of L2 word combinations

To the best of our knowledge, only two studies to date have looked at the acquisition of L2 collocations and other types of word combinations by adopting a phraseological framework. Nesselhauf (2003) evaluated the use of word combinations by L2 learners of English in free written production, and found the highest proportion of errors in collocations (79%, *take a picture*), compared to free combinations (23%, *want a car*) and idioms (23%, *sweeten the pill*). In particular, collocations with a medium degree of restriction showed the highest percentage of errors (33%), while the lowest proportion was found in collocations with a low degree of restriction (18%). The most frequent types of errors were the choice of a wrong verb (**carry out races* vs. *hold races*), the choice of a wrong noun (**close lacks* vs. *close gaps*), and the production of a wrong combination (*take notice* vs. *to notice*).

In a recent study, Gyllstad and Wolter (2016) compared the processing of collocations (*run a risk*) versus free combinations (*write a letter*). They employed a semantic judgment task, in which participants decided as quickly and accurately as possible whether or

not target word pairs were meaningful and natural English phrases. These authors found no significant differences between L1 and L2 speakers in terms of the retrieval of the overall meaning. Moreover, both groups responded faster to collocations and free combinations than to their respective controls. Further, higher frequency items were processed faster than lower frequency ones. A processing cost was found for collocations compared to free combinations in both groups of speakers. They argued that the processing cost observed for collocations might be due to the fact that collocations are less transparent than free combinations. Specifically, the verb in a collocation is often used in a delexical sense (the meaning of the verb does not contribute to the meaning of the combination). For a delexical verb (*make*), the authors claim, there is a processing competition between its concrete and abstract sense that can only be solved once the noun is processed. As authors argue, 'this sense resolution is believed to come with a processing cost' (p. 314). Gyllstad and Wolter (2016) concluded that semantic transparency affects the processing of lexical combinations, offering support to Howarth (1998).

Both Nesselhauf (2003) and Gyllstad and Wolter (2016) studies gave support to phraseological models and to the claim that collocations might be more challenging for learners compared to other types of word combinations (Henriksen, 2013). In addition to the above studies, a number of studies investigated the productive knowledge of Verb+Noun collocations in L2 learners using cloze and fill-in tasks (Bahns and Eldaw, 1993; Farghal and Obiedat, 1995), finding that learners were likely to produce collocations with a wrong collocate (*write a diary*, instead of *keep a diary*). In addition, Farghal and Obiedat (1995) claimed that synonymy was the most frequent strategy adopted by L2 learners (*fast colours* vs. *steady colours*).

In one of his early studies, Gyllstad (2005) investigated receptive knowledge of L2 Verb+Noun English collocations. Two lexical combinations were juxtaposed, one typical collocation and one pseudo-collocation, created by substituting the verb of the collocation with a synonym (*pay a visit* vs. *do a visit*). Participants decided which of the two forms was a typical collocation in English. Results showed that learners had not acquired correct use of collocations.

Finally, Sonbul (2015) investigated L1 and L2 speakers' sensitivity to collocational frequency using both offline (typicality rating task) and online (eye-tracking) measures. Adjective+Noun collocations (*fatal mistake*) were used with the adjective substituted with a synonym (a low frequency collocate of the noun, *awful mistake*, or a non-collocate of the noun, *extreme mistake*). Participants were asked to rate how typical collocations were in English. Typicality scores were positively affected by phrase frequency. Interestingly, average scores for non-collocations in both groups were above 3 on a 6-point scale. The author argued that collocations are more lexically flexible in comparison to other types of word combinations.

In conclusion, the studies reviewed above (Bahns and Eldaw, 1993; Farghal and Obiedat, 1995; Gyllstad, 2005; Sonbul, 2015) suggest that L2 learners are likely to prefer the use of a synonym over the typical element within a collocation. This implies that learners may operate on the open-choice principle rather than on the idiom principle (Sinclair, 1991), and that they do not perceive the degree of lexical fixedness of collocation. Overall, few studies have investigated L1 and L2 speakers' intuitions about the degree of lexical fixedness of collocation (see Gyllstad, 2005; Sonbul, 2015). Fewer still

have focused on L1 and L2 speakers' intuitions about various types of word combinations that differ in lexical fixedness and compositionality.

II The present study

Driven by the pertinent research in the area, we carried out two studies to investigate L2 learners' intuitions about word combinations and to explore whether or not they are affected by lexical fixedness. In keeping with tradition in the field, we employed a L1 group as a control group, our baseline. Our primary motivation was *not* to show that L2 intuitions may be different from L1 intuitions (this is not unexpected), but to demonstrate *how* they may be different, that is, how experience with language may modulate speakers' perceptions of word combinations and their key properties. To this aim, we adopted a phraseological approach and singled out three types of word combinations: free combinations, collocations and idioms. Word combinations were identified using a corpus-based method. We further considered whether compositionality affected speakers' intuitions about perceived restriction of word combinations, and explored the extent to which compositionality interacted with lexical fixedness in L1 and L2 intuitions about the use of word combinations.

III Study 1

Study 1 used a judgment task to investigate L2 learners' and L1 speakers' (control group) intuitions about Italian Verb+Noun word combinations. In particular, we wanted to investigate whether the three types of word combinations may be perceived differently in terms of lexical fixedness. Study 1 sought to answer the following questions:

1. How are word combinations perceived in terms of lexical fixedness?
2. Does the degree of lexical fixedness affect language users' intuitions?
3. Does compositionality affect language users' intuitions about the restriction of word combinations?

1   Participants

Twenty speakers of L1 Italian (13 females; age range: 24–37, Mean = 29.5, SD = 3.7) and 20 learners of L2 Italian (L1 English) took part in Study 1 (18 females; age range: 19–57, Mean = 25, SD = 11.2). L2 learners were full-time students studying Italian at a range of universities in New Zealand (Victoria University of Wellington, University of Auckland), the UK (University of Birmingham, Durham University, University of Edinburgh, University of Saint Andrews, University of Glasgow), and Italy (University of Pisa, University of Parma, University for Foreigners of Siena). L1 speakers of Italian were full-time students at the University of Pisa, University of Parma, and University for Foreigners of Siena. The present study was carried out in line with the ethical procedures at Victoria University of Wellington. L2 learners completed a language background questionnaire reporting their prior experience with the Italian language. They further rated their speaking, writing, listening and reading on a 5-point Likert scale (1 = very

Table 1. Language proficiency of second language (L2) speakers of Italian.

	Means	SD	Range
First contact with Italian (in years)	21.4	10.5	13–53
Time spent in Italy (in months)[a]	2.2	0.7	1–4
Time learning Italian (in months)	42.9	31.5	6–156
Speaking[b]	3.4	0.7	2–5
Writing[b]	3.9	0.5	3–5
Listening[b]	3.8	0.9	2–5
Reading[b]	3.9	0.9	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').

poor; 2 = weak; 3 = ok; 4 = good; 5 = excellent) and specified their level according to the Common European Framework of Reference for Languages, or CEFR (A2 = 2, B1 = 6, B2 = 25, C1 = 8, C2 = 1).[2] Table 1 summarizes L2 learners' experience with and knowledge of Italian.

2 Materials

Italian Verb+Noun free combinations and collocations were manually extracted from the ItTenTen corpus[3] (Jakubíček et al., 2013). In order to distinguish free combinations from collocations, we adopted a phraseological approach, using the criteria proposed by Ježek (2005) and Masini (2009). Following these criteria, we defined free combinations as word sequences that are not subject to lexical restriction and allow the substitution of their members (e.g. *guardare/vedere un film* 'to watch a movie'). Collocations were defined as frequent word combinations characterized by lexical fixedness allowing limited substitution of the constituents (e.g. *mantenere l'equilibrio* vs. **conservare l'equilibrio* 'to keep balance').

We preliminarily extracted 200 Verb+Noun combinations from the reference corpus and classified them as either free combinations or collocations, based on the criteria of lexical fixedness and compositionality. To identify collocations, we further used a frequency-/association-based approach. Mutual information (MI) and t-score are often used to identify significant collocations. We extracted collocations with a t-score of 2 or above, and MI score of 3 or above (Evert, 2008; Hunston, 2002).

From the preliminary group of word combinations, we selected 40 free combinations (*indovinare una risposta* 'to guess the answer') and 40 collocations (*sporgere denuncia* 'to press charges'). On the basis of their ItTenTen phrase frequencies, half of the items of each type were assigned to a high frequency band (20 free combinations and 20 collocations; phrase frequency range: 981–9113), and the other half to a low frequency band (20 free combinations and 20 collocations; phrase frequency range: 37–560).

In addition to collocations and free combinations, we selected Italian idioms using the databases created by Tabossi et al. (2011), which contains descriptive normative

Table 2. Summary of the means, the medians and the standard deviations of first language (L1) speakers' judgments of literality and lexical fixedness.

	Literality judgments			Lexical fixedness judgments		
	Mean	Median	SD	Mean	Median	SD
Free combinations	2.18	2.06	0.67	3	3.29	2.16
Collocations	2.52	2.31	0.81	4	3.81	2.18
Idioms	5.89	6.10	0.68	6	5.45	1.98

measures for 245 Italian idioms. Unlike free combinations and collocations, idioms are highly fixed and do not allow the substitution of their components with a synonym.

We selected 40 idioms (*rompere il ghiaccio* 'to break the ice') from our preliminary list. On the basis of the frequencies obtained from the ItTenTen corpus, idioms were assigned to two frequency bands: high frequency (20 items, phrase frequency range: 1191–6088) and low frequencies (20 items, phrase frequency range: 18–609).

Thus, in total, we selected 120 word pairings (40 free combinations, 40 collocations, 40 idioms). In order to ensure that the 120 lexical combinations differed in lexical fixedness and literality, we conducted two norming studies with 30 L1 speakers of Italian. In the first norming task, L1 speakers of Italian rated word combinations on a 7-point scale (1 = 'totally free', 7 = 'totally restricted'). We expected L1 speakers to judge free combinations as the least restricted word combinations, idioms as highly fixed word combinations, and collocations somewhere in-between. Word combinations were embedded in sentence contexts and presented in uppercase letters. In the second norming task, we aimed to test the literality of the word combinations, expecting that idioms would be perceived as non-literal, and collocation and free combination as literal. The 120 experimental items were presented in the same sentence contexts in uppercase letters. Thirty L1 speakers of Italian (not otherwise involved in the study) judged literality/figurativeness of word combinations on a 7-point scale (1 = 'totally literal', 7 = 'totally idiomatic'). The results of the norming studies showed that idioms were judged as the most lexically fixed and the least literal combinations, free combinations as the least lexically restricted and the most literal word combinations. Collocations appeared to be in-between free combinations and idioms (Table 2).

In Study 1, the 120 word combinations were embedded in sentence contexts. In order to investigate whether L1 and L2 speakers are sensitive to lexical fixedness, and if they judge as acceptable the substitution of the elements, we created another 120 sentences substituting the verb of the combinations with a synonym. Synonyms were extracted from the Italian section of MultiWordNet (Pianta et al., 2002) and were matched with the original word in frequency, MI and t-score. This resulted in 240 sentences: 120 containing original word combinations (with the original verb) and 120 containing variant word combinations (with the synonym of the verb).

Example 1:
Free combination

La nonna **spedisce** una lettera al nipote. (original)

La nonna **invia** una lettera al nipote. (variant)

'The grandmother posts/sends a letter to her grandson.'

Example 2:
Collocation

L'amico **ha mantenuto** la promessa. (original)

L'amico **ha conservato** la promessa. (variant)

'The friend kept/retained his promise.'

Example 3:
Idiom

Il professore **ha perso** il filo del discorso. (original)

Il professore **ha smarrito** il filo del discorso. (variant)

'The professor lost/misplaced the thread of the conversation.'

3   Computational indices of compositionality

In order to investigate lexical flexibility across the three types of combinations, we made use of computational indices of compositionality that rely on the degree of lexical flexibility of the construction under investigation (Senaldi et al., 2016, 2017). Distributional semantic indices of compositionality have been shown to possess psycholinguistic plausibility and to successfully predict phrase similarity in human judgments of acceptability (Mandera, 2017; Mitchell and Lapata, 2010) and in judgments of syntactic flexibility of idioms (Lebani et al., 2015).

To derive indices of compositionality for the target word combinations, we first used the variant-based distributional measures proposed by Senaldi et al. (2016, 2017). Senaldi et al. (2016) proposed a distributional algorithm for idiom type classification that was tested on a set of Italian Verb+Noun idiomatic and non-idiomatic expressions. For each construction, they generated a series of lexical variants by substituting the verb and the argument with semantically related words obtained distributionally, or manually selected from the Italian version of MultiWordNet (Pianta et al., 2002). The authors built a Distributional Semantic Model (DSM)[4] to represent the target expressions (idiomatic and literal) and their variants with vectors,[5] and to record their distributional association with linguistic contexts. Different measures were used by Senaldi et al. (2016) to compute the vector representations of the target expressions with those of their respective variants. In a follow-up study, Senaldi et al. (2017) successfully tested the same variant-based measures on a set of adjective−noun Italian idioms and non-idioms and compared them to the addition-based and multiplication-based indices that are traditionally used in the distributional literature on compositionality modelling (Mitchell and Lapata, 2010).

Senaldi et al.'s (2016, 2017) compositionality measures were applied to the dataset of this study and thus extended to collocations and free combinations. Following the same methodology, we built a DSM on the ItWac (Baroni et al., 2009) corpus using the top 30,000 content words, word combinations and their variants as target vectors and the 30,000 content words

Table 3. Summary of means, standard deviations and ranges for the two computational indices for the three types of word combinations.

	Additive index			Centroid		
	Mean	SD	Range	Mean	SD	Range
Free combinations	0.65	0.16	0.37−0.87	0.26	0.18	0.08−0.64
Collocations	0.64	0.12	0.20−0.77	0.26	0.17	0.03−0.46
Idioms	0.58	0.11	0.21−0.67	0.17	0.08	0.07−0.43

as contexts. The resulting matrix of co-occurrence was weighted with Positive Pointwise Mutual Information[6] (PPMI; Evert, 2008) and reduced to 300 dimensions via Singular Value Decompostion[7] (SVD; Deerwester et al., 1990). These parameters were set based on the previous parameter space explorations by Senaldi et al. (2016, 2017) and on common practice in the distributional literature (Kiela and Clark, 2014; Lapesa and Evert, 2014).

For each word combination and each synonymic variant, a lexical variant was created combing the verb of each word combination with its ten nearest neighbour nouns, and the noun of each word combination with its ten neighbour verbs. The semantic similarity of the target expressions (word combinations and their synonymic variants) was compared with cosine similarity.[8] After observing the correlational structure of our compositionality indices through hierarchical clustering, we decided to use centroid (the cosine similarity between the vector of a target expression and the centroid of the vectors of its variants) and the addition-based index (the cosine similarity between the vector of a target expression and the vector resulting from the sum of the vectors of the components of the target expression) as compositionality measures for the word combinations in our dataset.[9]

The rationale behind employing computational measures of semantic compositionality beside human ratings lies in the fact that such indices take advantage of corpus information that reveals the actual distributional behaviour of target expressions in general language use. On the one hand, the additive-based index indicates the extent to which the distributional behaviour of a given phrase in language use can be approximated by the summed distributional behaviour of its subparts. On the other hand, the centroid index grasps systematicity, which is tightly linked to semantic compositionality (Fodor and Lepore, 2002). Accordingly, if speakers can understand *spill the beans* used literally and *drop the peas*, they will also understand *spill the peas* and *drop the beans*. However, the same will not apply if we consider *spill the beans* as an idiomatic phrase. Non-compositionality thus goes hand in hand with lack of systematicity and our centroid index can model this in computational terms. A summary of the computational indices used for all the three types of word combinations can be found in Table 3.

4   Procedure

L2 and L1 speakers of Italian were asked to rate the 240 sentences on a 7-point Likert scale (1 = 'totally unacceptable', 7 = 'totally acceptable'). Because some of our

participants were L2 learners of Italian, we included an 'I don't know' option. The task was administered online through Figure Eight platform (https://appen.com/) with each participant being provided with a link. Item order was randomized. Detailed instructions were provided in both Italian and English; however, the task itself was in Italian. L1 speakers saw both the instructions and the task (without the 'I don't know' option) in Italian.

5   Analysis and results

We hypothesized that L1 speakers and L2 learners should judge the 'original version' of the lexical combinations as more acceptable than the 'variant version'. We expected the difference to be most prominent in idiomatic expressions, followed by collocations, with the original and the variant of free combinations being judged as equally acceptable. Finally, we wanted to investigate how L2 learners' intuitions may vary as a function of proficiency. We expected the accuracy of the judgments to increase in more proficient versus less proficient learners.

All the analyses were performed using mixed-effects modelling in R (Baayen et al., 2008). Two models were built in order to investigate L1 speakers' and L2 learners' intuitions. The models were built using R version 3.3.3 (2017-03-06) and the R packages lme4 (version 1.0-6; Bates et al., 2012), lsmeans (version 2.27-61), and lmerTest (version 2.0-6). In both models, we followed a step-by-step forward model selection procedure (Manning, 2007), based on the Akaike information criterion (AIC) values to select the most plausible model (Symonds and Moussalli, 2011; Wagenmakers and Farrell, 2004). We first checked for collinearity among item-related predictors. Item-related variables were not correlated. All continuous predictors were centred. We started with an empty model that included only participants and items as random effects, and added one predictor at a time by selecting the model with the lowest AIC value. We included predictors only if, when adding them, the AIC value decreased. We repeated the forward selection step procedure until no further predictors could be added. Then, we added one interaction term at a time by selecting the model with the lowest AIC value. We added interaction terms only if, when adding them, the AIC value decreased. We repeated the forward selection step procedure until no further interaction terms could be added.

The following predictors where included in the first model which used L1 judgements as the dependent variable: (1) type of combination (free combinations, collocations, idioms); (2) original/variant; (3) phrase frequency; (4) frequency band; (5) lexical fixedness (based on L1 speakers' judgments collected in the first norming task); (6) centroid; and (7) additive index. The frequencies of word combinations were logarithmically transformed. We further included the following interactions: (1) type of combination*original/variant; (2) type of combination*phrase frequency; (3) type of combination*frequency band; (4) original/variant*phrase frequency; (5) original/variant*frequency band; (6) additive index*lexical fixedness; (7) centroid*lexical fixedness; (8) additive index*original/variant; (9) lexical fixedness*original/variant; (10) centroid*original/variant. Participants and items were included as random effects, and L1 speakers' judgments as the dependent variable.

Table 4. Summary of means and standard deviations of participants' judgments of acceptability for each type of word combination.

	Original		Variant	
	Mean	SD	Mean	SD
First language (L1) speakers:				
Free combination	6.69	0.95	6.20	1.53
Collocation	6.50	1.22	4.81	2.42
Idiom	6.32	1.41	1.9	1.42
Second language (L2) speakers:				
Free combination	4.61	1.72	5	1.73
Collocation	5.26	1.66	5.10	1.71
Idiom	4.42	1.85	3.30	1.84

Table 5. Summary of the model for L1 speakers' judgments of word combinations.

	Estimate	SE	df	t value	Pr(>|t|)
(Intercept)	6.29	1.16	239.8	5.44	<.001
Phrase frequency	0.95	0.16	246.4	5.94	<.001
Frequency bands	0.18	0.14	236.7	1.29	.197
Type of combinations	1.89	0.44	243.1	4.28	<.001
Original/Variant	−0.36	0.73	258.6	−0.49	.618
Additive index	−4.22	1.66	237.3	−2.53	.011
Lexical Fixedness	0.55	0.26	237.1	−2.02	.044
Type of combination: Original/Variant	0.21	0.27	261.7	0.78	<.001
Phrase frequency: Type of combinations	−0.58	0.14	237.9	−4.13	<.001
Additive index: Lexical Fixedness	0.88	0.39	237.2	2.23	.026
Phrase frequency: Original/Variant	0.30	0.15	242.2	2.02	.044
Original/Variant: Lexical Fixedness	−0.26	0.16	240.8	−1.55	.122

In the second model, we used L2 learners' judgments as the dependent variable. The same predictors were included as in the first model (above), as well as proficiency[10] and the following interactions: (1) original/variant*proficiency; (2) type of combination*proficiency; (3) proficiency*phrase frequency; (4) proficiency*frequency bands; (5) proficiency*lexical fixedness; (6) proficiency*centroid; and (7) proficiency*additive index. Participants and items were included as random effects. The items responded to as 'I don't know' by L2 learners were excluded from the analysis (free combinations = 11%, idioms = 29.5%, collocations = 21.6%). The frequencies of word combinations were logarithmically transformed. Table 4 shows means and standard deviations of participants' judgments of acceptability for original and variant word combinations.
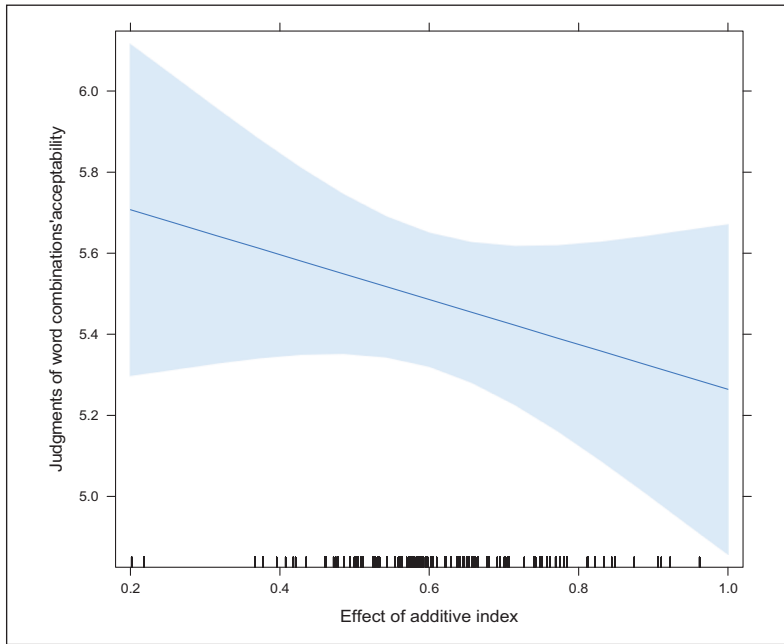
Figure 1. The negative effect of compositionality on L1 speakers' judgments.

6 L1 speakers

The final model, the main effects and their p-values can be found in Table 5. The following main effects were significant: type of combination, phrase frequency, additive index, lexical fixedness variables. Further, the following interactions were significant: type of combination*original/variant; type of combination*phrase frequency; original/variant*phrase frequency; lexical fixedness*additive index.

Given our research questions, we were mainly interested in the effect of compositionality and the interaction between the type of combination and original/variant. A negative effect of compositionality was found (Figure 1): as compositionality increased, the acceptability of L1 judgments decreased. L1 speakers judged lower compositionality combinations as more acceptable than higher compositionality combinations.

The speakers' judgments also varied as a function of both the type of combination and original/variant (Figure 2). L1 speakers judged the original combinations as more acceptable than their variants across all combination types. This is evident in the case of idiomatic expressions: original idioms received very high scores of acceptability while variant idioms received very low scores of acceptability. In addition, both original free combinations and their variants received high scores of acceptability, with original free combinations judged as only slightly more acceptable than their variant counterparts. Original collocations were judged as more acceptable than their variants.
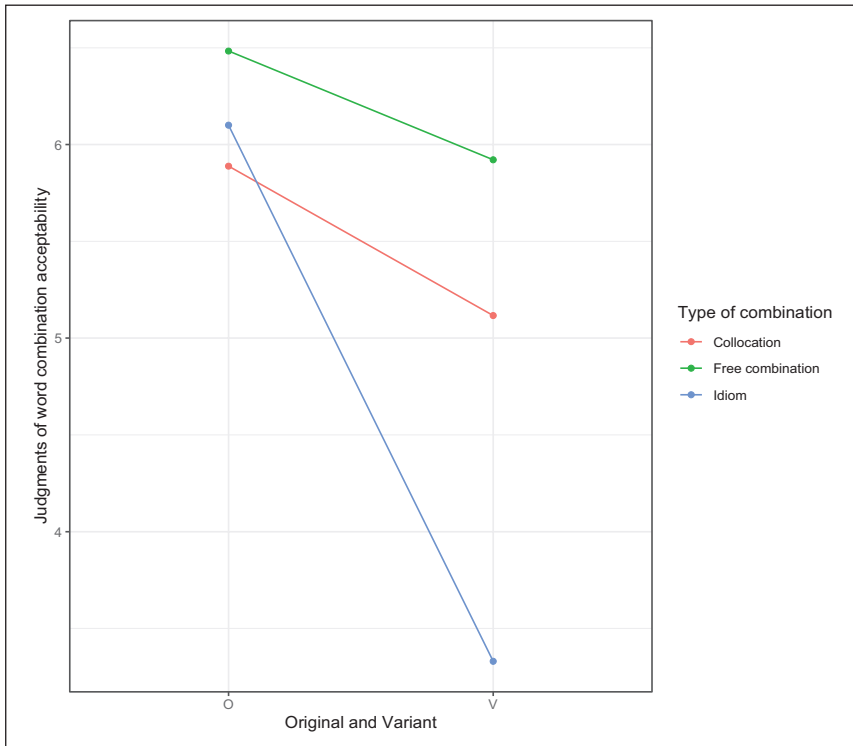
Figure 2. L1 speakers' judgments as a function of the type of combination and original/variant.

Table 6. L2 speakers' judgments of acceptability of the three combination types.

|	Estimate	SE	df	t value	Pr(>|t|)
Intercept	4.82	0.64	91.7	7.51	<.001
Type of combinations	0.12	0.36	2799.7	0.33	.003
Phrase frequency	0.21	0.06	232.9	3.38	<.001
Lexical fixedness	−0.13	0.07	234.5	−1.80	.072
Original/Variant	0.28	0.37	2771.2	0.75	.451
Proficiency	−0.05	0.15	35.2	−0.31	.756
Frequency bands	−0.48	0.21	2371.2	−2.26	.024
Type of combination: Original/Variant	0.30	0.54	2878.8	0.57	.023
Proficiency: Type of combination	0.28	0.11	3601.9	2.69	.008
Original/Variant: Proficiency	−0.01	0.10	3595.6	−0.15	.848
Proficiency: Frequency bands	0.16	0.05	3587.6	2.82	.004
Type of combination: Original/Variant: Proficiency	−0.20	0.15	3589.2	−1.38	.011
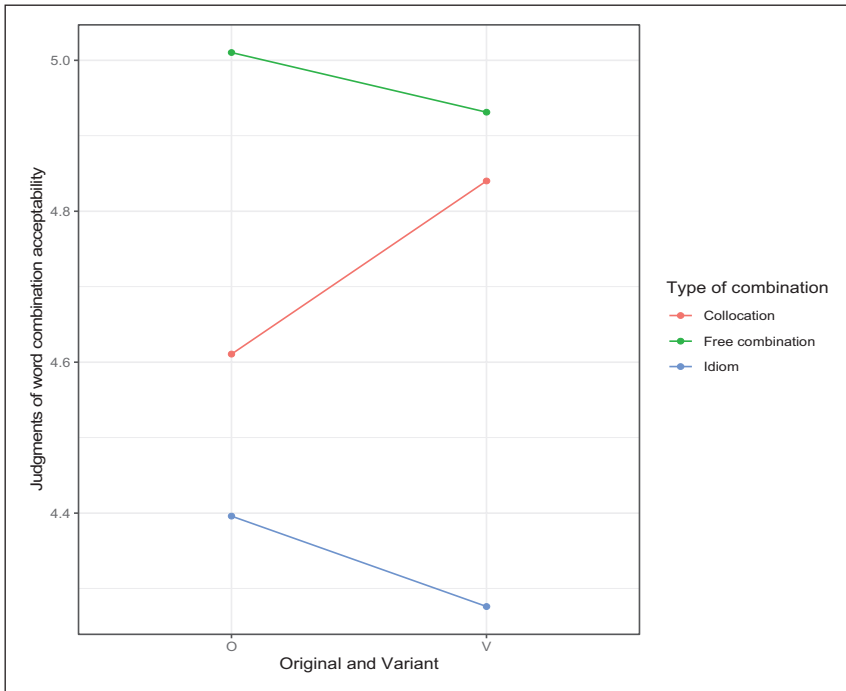
Figure 3. L2 speakers' judgements of original and variant combinations affected by the type of combination.

7 L2 learners

Table 6 shows the final model and the main effects. Three significant predictors were found: the type of combination, frequency band and phrase frequency. We also observed three significant two-way interactions: type of combination*original/variant, type of combination*proficiency, and proficiency*frequency band, as well as a significant three-way interaction between the type of combination, original/variant and proficiency.

Analyses of the L2 data revealed that acceptability judgments given to the original and variant combinations were affected by the type of combination (Figure 3). Original free combinations and idioms were judged as more acceptable than their variant counterparts. On the contrary, variant collocations were judged as more acceptable than original collocations.

Finally, the type of combination and the original/variant condition interacted with L2 proficiency and affected participants' judgments differently. Original and variant free combinations received relatively higher acceptability scores, and were judged overall as more acceptable than the other combination types (Figure 4). On the contrary, original idioms were judged as markedly more acceptable than their variant counterparts. Lower proficiency learners gave both original and variant idioms low scores of acceptability, while higher proficiency learners judged the original as more acceptable than the variant. Interestingly, all participants, across all proficiency bands, judged original collocations in a comparable way.
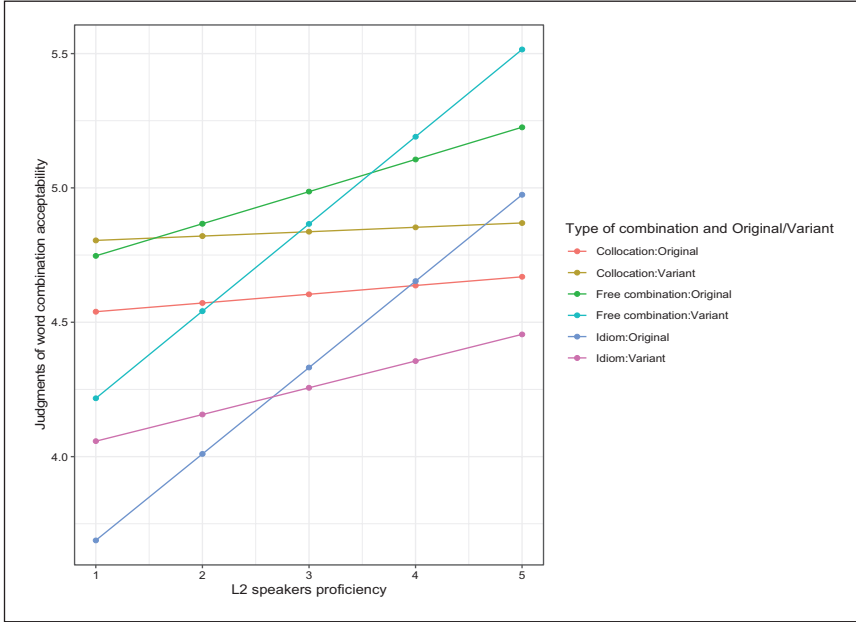
Figure 4. L2 speakers' judgments affected by the type of combination and the original/variant.

On the contrary, variant collocations were judged as more acceptable than original collocations by higher proficiency learners compared to lower proficiency learners.

IV Study 2

Study 2 used a multiple-choice test to investigate L2 learners' and L1 speakers' intuitions about the use of word combinations. In particular, Study 2 aimed at examining the role of lexical fixedness – that is, substitutability with a synonym – and compositionality in the use of word combinations. We wanted to answer the following questions:

1. How accurate are language users' intuitions about the use of word combinations?
2. How likely are language users to select a synonym of the verb rather than the original verb of a word combination?
3. Does compositionality affect speakers' use of word combinations?

1 Participants

Twenty speakers of L1 Italian (13 females; age range: 21–56; Mean = 30.7; SD = 8.8) and 22 learners of L2 Italian (L1 English) took part in Study 2 (17 females; age range:

Table 7. Language proficiency of L2 speakers of Italian.

	Means	SD	Range
First contact with Italian (in years)	22.3	12.8	6–70
Time spent in Italy (in months)[a]	2.2	0.7	1–4
Time learning Italian (in months)	40.7	30.1	6–120
Speaking[b]	3.1	1	1–5
Writing[b]	3.4	0.7	2–5
Listening[b]	3.4	0.7	1–5
Reading[b]	4	0.5	3–5

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").

10–70; Mean = 28; SD = 14.6). Participants in Study 2 were drawn from the same populations as in Study 1 (see above). Table 7 summarizes L2 learners' experience with and knowledge of Italian.

2 Materials

We used the same word combinations and sentences as in Study 1. We removed the verb from each sentence. Each sentence was presented with two possible verb options: the original verb (the 'original' response), and its synonym (the 'variant' response). Given that in free combinations either of the two verbs is acceptable and can be freely substituted, we added a 'both' option. Because some participants were L2 speakers of Italian, we also included an 'I don't know' option.

Example 4:
Free combination
La nonna _____ una lettera al nipote.
The grandmother _____ a letter to her grandson.
1. *spedire* ('posts'); 2. *inviare* ('sends'); 3. Both; 4. I don't know.
'The grandmother posts/sends a letter to her grandson.'

Example 5:
Collocation
L'amico ha _____ la promessa.
The friend _____ his promise.
1. *mantenuto* ('keep'); 2. *conservare* ('retain'); 3. Both; 4. I don't know.
'The friend kept/retained his promise.'

Example 6:
Idiom

Il professore ha _____ il filo del discorso.

The professor _____ the thread of the conversation.

1. *perso* ('lost'); 2. *smarrito* ('misplaced'); 3. Both; 4. I don't know.

'The professor lost/misplaced the thread of the conversation.'

3   Procedure

Participants chose the verb they thought was the most acceptable given the context. They could select 'both' if they felt both verbs were equally acceptable. The task was administered online through Figure Eight platform and each participant was provided with a link. Detailed instructions were provided in English. However, the task was in Italian. L1 speakers saw the instructions and the task (without the 'I don't know' option) in Italian.

4   Analysis and results

We hypothesized that 'both' should be most common in free combinations and least common in idioms, with collocations being in-between. Further, we hypothesized that 'original' responses should be more common than 'variant' responses in collocations and idioms.

5   L1 speakers

First, we looked at L1 speakers who selected 'original', 'variant' and 'both' responses. The proportion of the 'original' response was higher in the case of idioms (96%) compared with the 'variant' (1.8%) and 'both' (2.1%) responses.

 The proportion of the 'original' response (66.7%) was higher compared to the 'both' (22.5%) and 'variant' (10.7%) responses also in the case of collocations. Free combinations received high proportions of the 'original' (42.4%) and 'both' (45.1%) responses. The 'variant' response was selected more frequently in free combinations (12.5%) compared to idioms (1.8%) and collocations (10.7%; Figure 5).

 In order to explore what could predict the choice of the response ('original', 'variant', 'both'), we used three separate binomial logistic regressions. The models were built using R version 3.3.3 (2017-03-06) and the R packages lme4 (version 1.0-6; Bates et al., 2012). We first checked for collinearity among item-related variables. Predictors were not correlated. All continuous variables were centred.

 We followed the same procedure as in Study 1 (i.e. a step-by-step forward model selection procedure based on the Akaike information criterion/AIC values). The first model was built to investigate what could predict the 'original' response. The following predictors were included in the model: (1) type of combination; (2) phrase frequency of the combination with the original verb (phrase frequency 1); (3) phrase frequency of the combination with the variant verb (phrase frequency 2); (4) frequency band; (5) lexical fixedness; (6) centroid; and (7) additive index. We included in the model the following interactions: (1) type of combination*phrase frequency 1; (2) type of combination*phrase
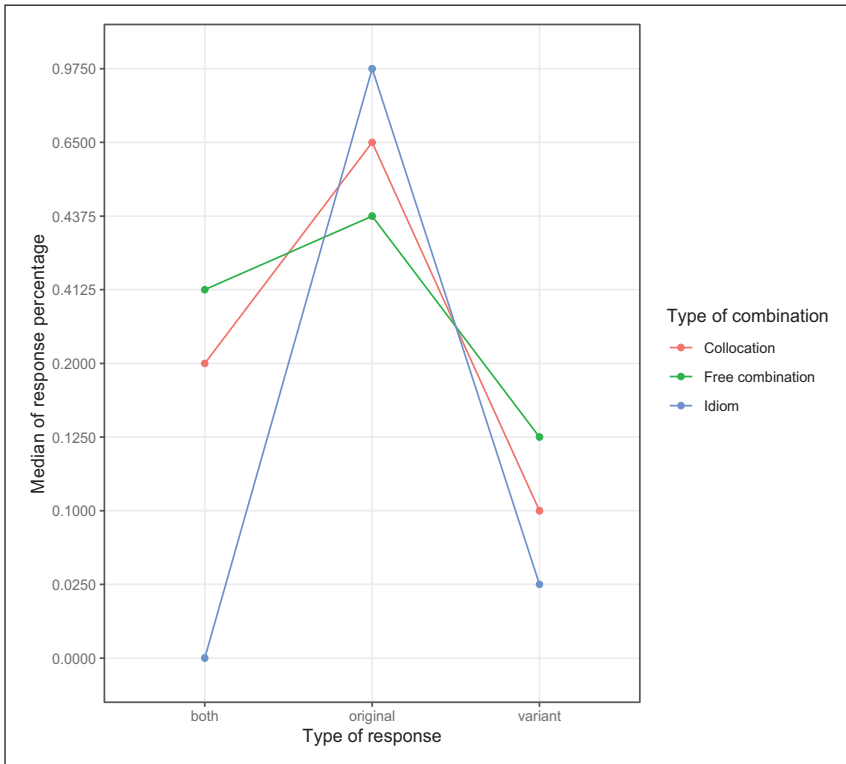
Figure 5. Distribution of L1 speakers' responses for the three types of word combinations.

frequency 2; (3) type of combination*frequency band; (4) phrase frequency 1*frequency band; (5) phrase frequency 2*frequency band; (6) additive index*lexical fixedness; and (7) centroid*lexical fixedness. Participants and items were included as random effects. The dependent variable indicated if L1 speakers selected the 'original' response (= 1) or not (= 0).

The final model and the main effects can be found in Table 8. The type of combination had a significant effect on selecting the 'original' response. Participants were more likely to select the 'original' response in collocations and idioms. Moreover, lexical fixedness and additive index significantly affected the selection of the 'original' response. The probability of selecting the 'original' response was higher in lower compositionality and higher lexical fixedness word combinations.

The second model investigated the 'variant' response. The same fixed and random effects as in the first model were included. The dependent variable indicated whether L1 speakers selected the 'variant' response (= 1) or not (= 0). We followed the same procedure as in the first model. Table 9 shows the final model and the main effects. The results indicate that the type of combination had a significant effect on the selection of 'variant' response.

Table 8. Summary of the first model (testing the "original" response in L1 speakers).

	Estimate	SE	Z value	Pr(>|z|)
(Intercept)[a]	7.97	3.05	2.61	.009
Type of combinations (idioms)	1.74	0.56	3.08	.002
Type of combinations (free combinations)	−0.65	0.33	−1.97	.048
Phrase Frequency 1	1.18	0.32	3.68	<.001
Phrase Frequency 2	−1.09	0.16	−6.85	<.001
Lexical fixedness	2.11	0.77	−2.71	.006
Additive index	−15.68	4.80	−3.26	.001
Lexical fixedness: Additive index	4.05	1.20	3.36	<.001

Note. Intercept levels: type of combinations = collocations.

Table 9. Summary of the second model (testing the "variant" response in L1 speakers).

	Estimate	SE	Z value	Pr(>|z|)
(Intercept)[a]	−21.88	4.19	−5.21	<.001
Type of combinations (idioms)	−2.24	0.72	−3.11	.002
Type of combinations (free combinations)	−0.11	0.39	−0.29	.769
Frequency bands	1.45	0.47	3.09	.002
Phrase Frequency 2	1.01	0.21	4.89	<.001
Lexical fixedness	−4.13	0.99	4.15	<.001
Additive index	24.24	5.98	4.05	<.001
Lexical fixedness: Additive index	−6.18	1.54	−4.01	<.001

Note. Intercept levels: type of combinations = collocations.

We found that lexical fixedness and compositionality significantly affected the probability of selecting the 'variant' response, and their interaction had a significant effect on the selection of the 'variant' response. The probability of selecting the 'variant' response was higher in higher compositionality and in less lexically restricted word combinations.

The third and final model explored the 'both' answer. The same fixed and random effects were included in this model as in the first and second models (see above). The dependent variable indicated whether participants selected the 'both' option (= 1) or not (= 0). We followed the same procedure as outlined above. The final model and the main effects can be found in Table 10. Again, the results indicate that the type of combination and phrase frequency 1 had a significant effect on the selection of 'both' response. Further, lexical fixedness significantly affected the selection of the 'both' response. The probability of selecting the 'both' response was higher in word combinations with a low degree of lexical fixedness.

6   L2 learners

The same analyses were conducted in order to investigate L2 speakers' responses to the three types of word combinations. First, we looked at the percentage of learners who

Table 10. Summary of the third model (testing the "both" response in L1 speakers including compositionality and lexical fixedness indices).

	Estimate	SE	Z value	Pr(>|z|)
(Intercept)[a]	1.25	1.01	1.23	.217
Type of combinations (idioms)	−1.07	0.53	−2.02	.0432
Type of combinations (free combinations)	0.84	0.30	2.80	.005
Phrase Frequency 1	−0.71	0.22	−3.13	.002
Lexical fixedness	0.66	0.22	−2.99	.003

Note. Intercept levels: type of combinations = collocations.

selected 'original', 'variant' and 'both' responses. Free combinations received higher proportions of the 'both' response (25.1%) compared to collocations (10.2%) and idioms (7.6%). The percentage of the 'original' response (47.1%) was higher than the percentage of 'both' (7.6%) and 'variant' (21.5%) in the case of idioms. Collocations received the 'variant' response (39.3%) more frequently than the 'original' response (32.3%). Figure 6 shows the distribution of learners' responses.

In order to explore what could predict the choice of the response ('original', 'variant' and 'both'), we used three separate binomial logistic regressions. The models were built using R version 3.3.3 (2017-03-06) and the R packages lme4 (version 1.0-6; Bates et al., 2012) and car (version 2.1-6, Fox and Weisberg, 2019).

L2 proficiency was treated as a continuous variable, with the four CEFR levels coded as follows: A2 = 1, B1 = 2, B2 = 3, C1 = 4. The items responded to as 'I don't know' were excluded from the analysis (free combinations 7.4%, idioms 23.8%, collocations 18.2%). The first model was built to investigate what could predict the 'original' response. We included the same predictors as those included in L1 speakers' analysis (above), as well as proficiency and the following interactions: (1) proficiency*phrase frequency 1; (2) proficiency*phrase frequency 2; (3) proficiency*frequency band; (4) proficiency*additive index; (5) proficiency*lexical fixedness; and (6) proficiency*centroid. Participants and items were included as random effects. The dependent variable indicated if learners selected the 'original' response (= 1) or not (= 0).

We proceeded with a step-by-step forward selection procedure to select the most plausible model (based on AIC values). We followed the same procedure as described above.

The final model and the main effects can be found in Table 11. Type of combination significantly affected the selection of the 'original' response. L2 learners were more likely to select the 'original' response in idioms and in free combinations compared to collocations. Results suggest that L2 speakers were less likely to select the 'original' response in collocations.

The second model investigated the 'variant' response. The same fixed and random effects as in the first model were included. The dependent variable indicated whether L2 speakers selected the 'variant' response (= 1) or not (= 0). We followed the same procedure as in the first model. We found a significant effect of the type of combination and proficiency. The final model and the main effects can be found in Table 12. We found a
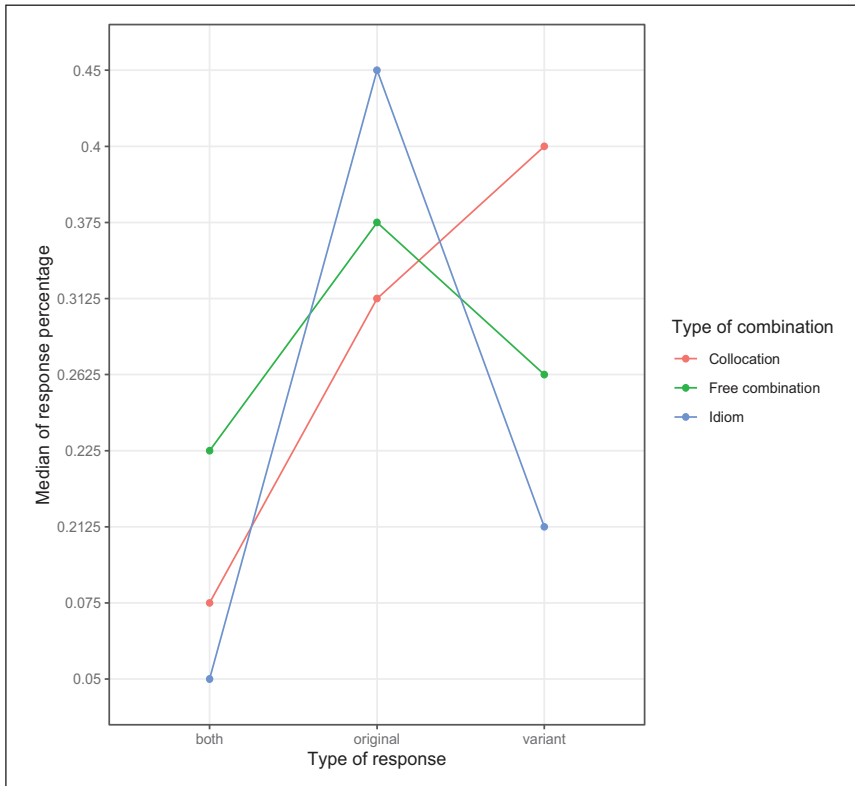
Figure 6. Distribution of L2 speakers' responses for the three types of word combinations.

Table 11. Summary of the first model (testing the "original" response in L2 speakers).

	Estimate	SE	Z value	Pr(>|z|)
(Intercept)[a]	−0.70	0.51	−1.36	.171
Type of combinations (idioms)	0.76	0.27	2.81	.004
Type of combinations (free combinations)	0.28	0.28	1.14	.251
Phrase frequency 2	0.02	0.18	1.15	.247
Proficiency	0.23	0.15	1.49	.135
Phrase Frequency 2: Proficiency	−0.16	0.05	−3.06	.002

Note. Intercept levels: type of combinations = collocations.

significant effect of the type of combinations: learners were more likely to select the 'variant' response in collocations compared to idioms and free combinations. Further, they were less likely to select the 'variant' response in free combinations and idioms compared to collocations.

Table 12. Summary of the second model (testing the "variant" response in L2 speakers).

	Estimate	SE	Z value	Pr(>|z|)
(Intercept)[a]	0.51	0.28	−1.79	.072
Type of combinations (idioms)	−0.79	0.31	−2.58	.009
Type of combinations (free combinations)	−1.09	0.28	−3.89	<.001
Phrase Frequency 2	0.23	0.11	2.17	.029

Note. Intercept levels: type of combinations = collocations.

Table 13. Summary of the third model (testing the "both" response in L2 speakers).

	Estimate	SE	Z value	Pr(>|z|)
(Intercept)[a]	−2.50	1.25	−1.99	.045
Type of combinations (idioms)	0.33	0.37	0.88	.376
Type of combinations (free combinations)	0.88	0.25	3.48	<.001
Lexical Fixedness	−0.277	0.16	−1.63	.102
Phrase frequency 2	−0.35	0.23	−1.49	.136
Frequency bands	0.55	0.21	2.53	.011
Proficiency	0.19	0.32	0.59	.554
Phrase frequency 2: Proficiency	0.16	0.07	2.37	.017

Note. Intercept levels: type of combinations = collocations.

Finally, we built a third model to explore the 'both' answer. The same fixed and random effects were included in this model as in the first and second models (see above). The dependent variable indicated whether participants selected the 'both' option (= 1) or not (= 0). We followed the same procedure as outlined above. Table 13 shows the final model and the main effect. Type of combination significantly affected the selection of 'both' response. Learners were more likely to choose the 'both' option in idioms and free combinations compared to collocations.

The three binomial logistic regressions showed that the type of combination affected the selection of the three responses. The learners were more likely to select the 'both' response in free combinations, the 'original' response in idioms, and the 'variant' response in collocations. However, we did not find any effect of compositionality or lexical fixedness on the selection of the three responses.

V General discussion

In this study, we sought to investigate language users' intuitions about three different types of word combinations: free combinations collocations and idioms. The present research had three aims. First, we wanted to investigate L1 speakers' and L2 learners' perceptions of word combinations in terms of lexical fixedness. Second, we sought further confirmation to the tenet that collocations are challenging for L2 learners, as

previously suggested in the literature (Laufer and Waldman, 2011; Wray, 2002). Finally, we wanted to establish whether the degree of compositionality of the target combinations affects language users' intuitions.

Two studies were carried out. In Study 1, Italian Verb+Noun combinations were embedded in sentence contexts. The combinations appeared either with the original verb, or with a synonym. L1 speakers and L2 learners rated sentence acceptability in both conditions. In Study 2, the verb was removed from each sentence while participants chose the verb (from the options provided) they felt was the most acceptable completion. The two studies employed different experimental designs and tasks (acceptability judgment and multiple-choice task, respectively); yet they showed a comparable pattern of results.

First, our analysis showed that idioms, collocations and free combinations were perceived differently in terms of their lexical fixedness by all participants, with idioms rated as least flexible of the three combination types. However, some notable differences emerged between L1 speakers and L2 learners. Namely, while collocations were rated as less flexible than free combinations by L1 speakers (in line with phraseological models), they were rated as more flexible than free combinations by L2 learners.

Specifically, L1 speakers judged the use of a synonym as more acceptable in free combinations than in collocations and idioms. They likewise selected the variant and the 'both' option more frequently in free combinations compared to collocations and idioms. L2 learners differed from L1 speakers in how they judged the substitution of the verb in collocations. Indeed, the substitution of the verb was judged as more acceptable in free combinations and collocations compared to idioms. Further, L2 learners were more likely to select the variant option in collocations and free combinations than in idioms.

As expected, the intuitions of our control (L1) participants were found in line with the lexical flexibility-based distinction between the three types of word combinations posited by phraseological models (Cowie, 1981; Howarth, 1998) and, in particular, by Ježek's (2005) and Masini's (2009) classifications. Phraseological models place collocations in-between free combinations and idioms, predicting synonym or near-synonym substitution to be impossible in some cases (*destare sospetto* 'to arouse suspicion', **svegliare sospetto* 'to wake suspicion'), and more acceptable in others (*disfare/svuotare la valigia* 'to unpack/empty the luggage'). On the contrary, the flexibility of idioms' components is highly restricted. The middle-ground position of collocations was evident in the intuitions of L1 speakers, who judged free combinations as the least fixed word combinations and idioms as the most fixed ones, and rated the variant form of collocations as acceptable only in certain cases. Our results are in contrast to those reported by Sonbul (2015), who found that L1 speakers gave high typicality judgments to variant collocations. Sonbul (2015) argued that the high ratings given to non-collocations 'emphasized the fluidness' of collocations (p. 13) relative to other word combination types (binomials, idioms), with language users likely to accept modifications of typical collocations.

The second aim of the present study was to offer support to the tenet that collocations are challenging for L2 learners. Indeed, L2 learners' intuitions were not found to be

accurate with respect to collocations. Both low and high proficiency learners judged their variant forms as more acceptable than the original ones. In line with earlier research, the results of the present study confirm that collocations (Verb+Noun collocations, in particular) are challenging for learners who often produce deviant collocations, substituting the verb in a collocation with a synonym (Bahns and Eldaw, 1993; Farghal and Obiedat, 1995; Gyllstad, 2005; Laufer and Waldman, 2011; Nesselhauf, 2003). Such difficulty might be due to several factors. First, although collocations are often semantically compositional and transparent (Howarth, 1998), the verb is often used in a figurative sense (*keep a promise* vs. *keep some food in the fridge*) (Gyllstad and Wolter, 2016). For instance, in the Italian collocation *accendere un mutuo* ('to take out a mortgage') the verb *accendere* ('to light') is used in a figurative sense meaning *aprire* ('to open'). If learners are not familiar with the collocation *accendere un mutuo*, they may not retrieve the figurative sense of the verb and likely to substitute it with a synonym (e.g. *iniziare un mutuo* 'to start a mortgage') producing a more literal and less idiomatic collocation than the original one. Further, Verb+Noun collocations present a higher degree of variability with respect to other types of collocations, such as Adjective+Noun / Noun+Adjective collocations (Peters, 2016; Wolter and Yamashita, 2015). For example, the verb in a collocation can occur in different forms (*She keeps a promise, She has kept a promise, She wants to keep a promise*). Further, Verb+Noun collocations vary in terms of presence or absence of determiners (*to make a mistake* vs. *to waste time*). In addition, Verb+Noun collocations may occur in their non-adjacent form (Vilkaitė, 2016; Vilkaitė and Schmitt, 2019), with an intervening modifier (*to waste time to waste too much time*) or another element (*to lend a hand to lend a helping hand*). Such modification is also common in Italian (*perdere tempo* 'to waste time', *perdere molto tempo* 'to waste a lot of time'; *dare aiuto* 'give/offer help', *dare un grosso aiuto* 'give/offer a lot of help').

L2 learners may find such variability challenging and confusing at best, meaning that the two content words within a collocation may not be as strongly associated, or linked, in a learner's mind as components of a highly fixed phrase. As a result, a learner may perceive a collocation as less fixed than it really is. Further, since many collocations are semantically compositional and are made of highly frequent words, it is likely that learners may not deem collocations as problematic elements in their input (Schmidt, 1994). That is, learners may not perceive collocations as prefabricated patterns, processing them, instead, word-by-word.

Third, we considered the effect of semantic compositionality on L1 speakers' and L2 learners' intuitions. From a methodological standpoint, we specifically aimed at assessing whether corpus-based data on semantic compositionality can predict human judgments of lexical fixedness. We experimented with two different computational measures of compositionality. The first, inspired by Senaldi et al. (2016, 2017), compares the distributional similarity of a target expression with that of its lexical variants and provides a computational approximation of how much altering the lexical choice in a word combination results in a meaning shift. The second addition-based index is traditionally employed in the distributional literature on compositionality (Krčmář et al., 2013; Mitchell and Lapata, 2010), and captures how similar the distributional behaviour of a word combination is to the sum of the distributional behaviours of the individual

components. Only the second additive index turned out to be significant in our models. Interestingly, the more the distribution of a given combination differed from that of its components (the less compositional the expression), the more language users were likely to rate its use as acceptable, probably by virtue of the fact that non-compositional combinations have a higher degree of conventionality. However, L2 speakers differed from L1 speakers, in that learners judged the variant form of a collocation as more acceptable than its conventional form. The preference of a synonym rather than the typical element of a collocation aligns well with the open choice principle (Sinclair, 1991).

Our results suggest that while L1 speakers appear to operate on the idiom principle, L2 speakers tend to operate on the open choice principle (Sinclair, 1991). The open choice principle is 'a way of seeing language text as the result of a very large number of complex choices' (Sinclair, 1991:110), with the only restriction being grammaticality and each slot being able to be filled with any word. On the contrary, according to the idiom principle 'a language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices' (Sinclair, 1991: 110). According to Sinclair (1991), proficient language users prefer an idiom choice principle to an open-choice principle in linguistic production, going for semi-preconstructed phrases rather than *ex-novo* compositional and combinatorial expressions. A decade earlier, Pawley and Syder (1983) put forward a similar distinction, claiming that speakers have access to both lexical items (open-choice principle) and memorized chunks (idiom principle). Language users are able to select conventional ways of expression among a wide range of grammatically acceptable phrases, some of which include grammatical but 'non-native-like or highly marked usages' (p. 191).

Similar results were reported in Singleton et al. (2007), who investigated the production of Intensifier+Adjective collocations where the intensifier displayed different degrees of specificity/generality (*very important*) by L1 speakers and advanced learners of L2 English. The authors found L1 speakers to be more likely to select specific intensifiers, and L2 speakers to be more likely to select general intensifiers, arguing that L1 speakers' production operates more on the idiom principle, whereas L2 learners' production is mostly based on an open-choice principle.

Our findings suggest that advanced learners, unlike lower proficiency learners, are aware of the lexical fixedness of idioms. This might be due to the fact that idioms are probably more salient in the input (Ellis, 2017), in that their restricted lexical choice is linked to the conveyance of a non-compositional meaning that often results in a literally meaningless combination. Contrary to idioms, many collocations are semantically compositional in spite of their lexical fixedness, so it is likely that learners with little exposure to L2 may not perceive the lexeme choice in collocations as restricted as in idioms.

Collocations, indeed, were found to be challenging for L2 learners. This might be due to limited exposure to the L2. Previous studies showed that encountering collocations frequently in L2 input strengthens the links between collocations and their conceptual representations (Gyllstad, 2007; Webb et al., 2013; Yamashita and Jiang, 2010). If learners are not sufficiently exposed to a L2 (as the learners who took part in our studies), the associative links between the elements of a collocation cannot be strengthened (Durrant and Schmitt, 2010).

Frequency of exposure to collocations in L2 input may affect their acquisition. In Ellis's (2001) model of collocation learning in a L1, collocations are the product of a 'chunking mechanism'. When two or more frequently co-occurring words are encountered, they are stored as a unique 'chunk'. A similar mechanism can be hypothesized to take place in L2 contexts, if L2 learners are surrounded by a great amount of linguistic evidence (Ellis, 2001).

Similar to Ellis (2001), Hoey (2005) claims that there is no difference in the mechanisms of L1 versus L2 acquisition. Collocations are acquired through exposure to language; language users store information about each and every encounter as well as the context of usage. According to Hoey (2005), the only difference between L1 and L2 acquisition of collocations is the amount of exposure. L1 learners acquire collocations inductively as they are exposed to huge amounts of language. If L2 learners are exposed to large amounts of the target language, they are likely to acquire collocations in a way similar to L1 speakers.

Indeed, the difference in the accuracy between L1 and L2 speakers' intuitions lies in the fact that L2 speakers did not have the same amount of input as L1 speakers. L1 speakers are constantly surrounded by linguistic evidence and are thus sensitive to the degree of lexical fixedness of collocations. On the contrary, L2 speakers are not exposed to large amounts of L2 and may not be aware of the lexical restriction of collocations. It is thus likely that learners with little exposure to L2 may operate on an open-choice principle. A greater amount of exposure to L2 may make learners more aware of and attuned to conventionality and lexical restriction of collocation.

In conclusion, the present study provides new insights into how L1 and L2 speakers perceive lexical flexibility of different kinds of word combinations, such as free combinations, collocations and idioms. We employed two offline tasks – acceptability rating and lexical choice – to obtain a more complete picture of the way in which L1 and L2 speakers may perceive lexical flexibility of the three types of combinations. In addition, our behavioural data were in part modelled with the help of computational indices of semantic compositionality, to see if such corpus-based evidence possesses psycholinguistic reality. Both behavioural data and computational indices were found to be powerful methods for investigating lexical properties of word combinations.

In sum, the findings reported in the present exploration suggest that L1 and L2 speakers perceive lexical flexibility in word combinations differently, with L1 speakers judging non-compositional combinations as more acceptable than compositional combinations, and L2 speakers having a harder time recognizing lexical restriction of collocations.

Acknowledgements

Author Contributions

Declaration of Conflicting Interests

Funding

ORCID iD

Irene Fioravanti [iD] https://orcid.org/0000-0001-5182-9394

Notes

1. Idioms have traditionally been defined as formally frozen and semantically empty non-compositional word combinations (Bobrow and Bell, 1973; Swinney and Cutler, 1979), although psycholinguistic and corpus-based evidence has revealed that formal variation is possible for semantically decomposable idioms (Gibbs and Nayak, 1989).
2. They reported the proficiency level they were assigned to at the end of the Italian course.
3. The ItTenTen corpus is a part of the TenTen corpus family which is a set of the web corpora with a target size $10+$ billion words.
4. Distributional Semantics (DS) is a usage-based approach to semantics, according to which linguistic context plays a role in characterizing and shaping the meaning of words and expressions. Semantic similarity is central to DS theory and is defined as a function of the degree to which they occur in similar contexts. DSMs make use of corpora and statistical methods to model and compare context-based representations of words and longer stretches of language. Contextual representations are conceived of as a combination of all the contexts in which the target expression is encountered.
5. Distributional vectors encode the co-occurrence statistics of the target expressions with specific contextual features.
6. It assesses whether two elements co-occur with higher probability than expected by chance alone, but then sets all negative results to zero.
7. It is commonly used in order to account for data sparseness and noisy data.
8. It is the most frequently used vector similarity measure, which computes the cosine of the angle between two vectors whose value spans between 0 and 1. If the two vectors are geometrically aligned, the angle between them is 0° and the cosine is 1 (highest semantic similarity). On the contrary, if the two vectors are perpendicular, the angle between them is 90° and the cosine is 0 (lowest semantic similarity).
9. Hierarchical clustering algorithm on compositionality scores and on compositionality and language users' judgments was conducted to explore the correlational structure to understand which predictors were more correlated with each other.

10. Learners' proficiency was treated as a continuous variable, with the five types of levels (according to CEFR) coded as follows: A2 = 1, B1 = 2, B2 = 3, C1 = 4, C2 = 5.

References

Baayen RH, Davidson DJ, and Bates DM (2008) Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59: 390–412.

Bahns J and Eldaw M (1993) Should we teach EFL students collocations? *System* 21: 101–14.

Baroni M, Bernardini S, Ferraresi A, and Zanchetta E (2009) The wacky wide web: A collection of very large linguistically processed web-crawled corpora. *Language Resources and Evaluation* 43: 209–26.

Bates DM, Maechler M, and Bolker B (2012) *lme4: Linear mixed-effects models using Eigen and S4: R package version 1.0-6*. Available at: https://cran.r-project.org/web/packages/lme4/index.html (accessed June 2020).

Bobrow SA and Bell SM (1973) On catching on to idiomatic expressions. *Memory and Cognition* 1: 343–46.

Carrol G and 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: 299–317.

Cowie AP (1981) The treatment of collocations and idioms in learners' dictionaries. *Applied Linguistics* 2: 223–35.

Cowie AP (1994) Phraseology. In Asher RE (ed.) *The encyclopedia of language and linguistics*. Oxford: Oxford University Press, pp. 3168–71.

Cowie AP (1998) *Phraseology: Theory, analysis and applications*. Oxford: Oxford University Press.

Deerwester S, Dumais ST, Furnas GW, Landauer TK, and Harshman R (1990) Indexing by latent semantic analysis. *Journal of the American Society for Information Science* 41: 391–407.

Durrant P and Schmitt N (2009) To what extent do native and non-native writers make use of collocations? *International Review of Applied Linguistics* 47: 157–177.

Durrant P and Schmitt N (2010) Adult learners' retention of collocations from exposure. *Second Language Research* 26: 163–88.

Ellis NC (2001) Memory for language. In Robinson P (ed.) *Cognition and second language instruction*. Cambridge: Cambridge University Press.

Ellis NC (2017) Salience. In: Hundt M, Mollin S, and Pfenninger S (eds) *The changing English language: Psycholinguistics perspectives*. Cambridge: Cambridge University Press, pp. 71–92.

Evert S (2008) Corpora and collocations. In: Ludeling A and Kyto M (eds) *Corpus linguistics: An international handbook*. Berlin: Mouton de Gruyter, pp. 1212–48.

Farghal M and Obiedat H (1995) Collocations: A neglected variable in EFL. *International Journal of Applied Linguistics* 28: 313–31.

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

Fodor J and Lepore E (2002) *The compositionality papers*. Oxford: Oxford University Press.

Fox J and Weisberg S (2019) *An R Companion to Applied Regression*. 3rd edition. Thousand Oaks, CA: SAGE.

Gibbs RW and Nayak NP (1989) Psycholinguistic studies on the syntactic behaviour of idioms. *Cognitive Psychology* 21: 100–38.

Gyllstad H (2005) Words that go together well: Developing test formats for measuring learner knowledge of English collocations. In: Heinat F and Klingval E (eds) *The Department of English in Lund: Working Papers in Linguistics* 5: 1–31.

Gyllstad H (2007) *Testing English collocations: Developing receptive tests for use with advanced Swedish learners*. Sprak-och litteraturcentrum: Lunds universitet.

Gyllstad H and Wolter B (2016) Collocational processing in light of the phraseological continuum model: Does semantic transparency matter? *Language Learning* 66: 296–323.

Hausmann FJ (1989) Le dictionnaire de collocations [Dictionary of collocations]. In: Hausmann FJ, Wiegand HE, and Zgusta L (eds) *Wörterbücher, dictionaries, dictionnaires: Ein internationales Handbuch zur Lexikographie*. Berlin: de Gruyter, pp. 1010–19.

Henriksen B (2013) Research on L2 learners' collocational competence and development: A progress report. In: Bardel C, Lindqvist C, and Laufer B (eds) *L2 vocabulary acquisition, knowledge and use: New perspectives on assessment and corpus analysis*. EuroSLA Monograph Series 2. Colchester: EuroSLA, pp. 29–56.

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

Howarth P (1998) Phraseology and second language proficiency. *Applied Linguistics* 19: 24–44.

Hunston S (2002) *Corpora in applied linguistics*. Cambridge: Cambridge University Press.

Jakubíček M, Kilgarriff A, Kovár V, Rychlý P, and Suchomel V (2013) *The TenTen corpus family*. *Proceedings of the 7th International Corpus Linguistics Conference*, pp. 125–27.

Ježek E (2005) *Lessico: Classi di parole, strutture, combinazioni* [Lexicon: Classes of words, structures and combinations]. Bologna: Il Mulino.

Kiela D and Clark S (2014) *A systematic study of semantic vector space model parameters*. In: *Proceedings of the 2nd workshop on continuous vector space models and their compositionality*. Association of Computational Linguistics, pp. 21–30.

Krčmář L, Ježek K, and Pecina P (2013) *Determining compositionality of expressions using various word space models and measures*. In: Allauzen A et al. (eds) *Proceedings of the Workshop on Continuous Vector Space Models and their Compositionality*. Association of Computational Linguistics, pp. 64–73.

Lapesa G and Evert S (2014) A large scale evaluation of distributional semantic models: Parameters, interactions and model selection. *Transactions of the Association for Computational Linguistics* 2: 531–45.

Laufer B and Waldman T (2011) Verb–noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning* 61: 647–72.

Lebani G, Senaldi MSG, and Lenci A (2015) *Modeling idiom variability with entropy and distributional semantics*. In: Wahle J et al. (eds) *Proceedings of the 6th Conference on Quantitative Investigations in Theoretical Linguistics*.

Mandera P, Keuleers E, and Brysbaert M (2017) Explaining human performance in psycholinguistic tasks with models of semantic similarity based on prediction and counting: A review and empirical validation. *Journal of Memory and Language* 92: 57–78.

Manning C (2007) Generalized linear mixed models. Unpublished course handout Stanford University, Stanford, CA, USA.

Masini F (2009) Combinazioni di parole e parole sintagmatiche [Combinations of words and syntagmatic words]. In: Lombardi Vallauri E and Mereu L (eds) *Spazi linguistic: Studi in onore di Raffaele Simone* [Linguistics spaces: Studies in honor of Raffaele Simone]. Roma: Bulzoni, pp. 191–209.

Mel'čuk I (1998) Collocations and lexical functions. In Cowie AP (ed.) *Phraseology: Theory, analysis and applications*. Oxford: Oxford University Press, pp. 23–53.

Mitchell J and Lapata M (2010) Composition in distributional models of semantics. *Cognitive Science* 34: 1388–1429.

Nesselhauf N (2003) The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics* 24: 223–42.

Pawley A and Syder FH (1983) Two puzzles for linguistic theory: Nativelike selection and native-like fluency. In: Richards JC and Schmidt RW (eds) *Language and communication*. London: Longman, pp. 191–225.

Peters E (2016) The learning burden of collocations: The role of interlexical and intralexical factors. *Language Teaching Research* 20: 113–38.

Pianta E, Bentivogli L, and Girardi C (2002) *MultiWordNet: Developing and aligned multilingual database*. In: *Proceedings of the First International Conference on Global WordNet*. Global WordNet Association, pp. 293–302.

Schmidt R (1994) Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review* 11: 11–26.

Senaldi MSG, Lebani G, and Lenci A (2016) *Lexical variability and compositionality: Investigating idiomaticity with distributional semantics model*. In: Calzolari N, et al. (eds) *Proceedings of the Tenth International Conference on Language Resources and Evaluation*. European Language Resources Association, pp. 21–31.

Senaldi MSG, Lebani G, and Lenci A (2017) Determining the compositionality of noun–adjective pairs with lexical variants and distributional semantics. *Italian Journal of Computational Linguistics* 3: 43–58.

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

Singleton D, Lesniewska J, and Witalisz E (2007) Open choice versus the idiom principle in L2 lexical usage. In: Pawlak M (ed) *Exploring focus on form in language teaching*. Studies in Pedagogy and Fine Arts. Poznan-Kalisz: Adam Mickiewicz University, pp. 207–22.

Siyanova A and Schmitt N (2008) L2 learner production and processing of collocation: A multi-study perspective. *The Canadian Modern Language Review* 64: 429–58.

Siyanova-Chanturia A, Conklin K, and Schmitt N (2011) Adding more fuel to the fire: An eye-tracking study of idiom-processing by native and non-native speakers. *Second Language Research* 27: 251–72.

Sonbul S (2015) Fatal mistake, awful mistake, or extreme mistake? Frequency effects on off-line/on-line collocational processing. *Bilingualism: Language and Cognition* 18: 419–37.

Swinney DA and Cutler A (1979) The access and processing of idiomatic expression. *Journal of Verbal Learning and Verbal Behavior* 18: 523–34.

Symonds MRE and 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: 13–21.

Tabossi P, Arduino L, and Fanari R (2011) Descriptive norms for 245 Italian idiomatic expressions. *Behave Res* 43: 110–23.

Underwood G, Schmitt N, and Galpin A (2004) The eyes have it: An eye-movement study into processing of formulaic sequences. In: Schmitt N (ed.) *Formulaic sequences*. Amsterdam: John Benjamins, pp. 153–72.

Vilkaitė L (2016) Are non-adjacent collocations processed faster? *Journal of Experimental Psychology: Learning, Memory, and Cognition* 42: 1632–42.

Vilkaitė L and Schmitt N (2019) Reading collocations in an L2: Do collocation processing benefits extend to non-adjacent collocations? *Applied Linguistics* 40: 329–54.

Wagenmakers EJ and Farrell S (2004) AIC model selection using Akaike weights. *Psychonomic Bulletin and Review* 11: 192–96.

Webb S, Newton J, and Chang A (2013) Incidental learning of collocation. *Language Learning* 63: 91–120.

Wolter B and Gyllstad H (2013) Frequency of input and L2 collocational processing. *Studies in Second Language Acquisition* 35: 451–82.

Wolter B and Yamashita J (2015) Processing collocations in a second language: A case of first language activation? *Applied Psycholinguistic* 36: 1193–1221.

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

Wray A (2002) *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.

Yamashita J and Jiang N (2010) L1 influence on the Acquisition of L2 Collocations: Japanese ESL Users and EFL Learners Acquiring English Collocations. *TESOL Quarterly* 44: 647–68.