

ARTIGOS

The use of collocations across proficiency levels: a literature review

Larissa Goulart¹ 

¹ Northern Arizona University, Flagstaff, AZ, Estados Unidos da América.

ABSTRACT:

This literature review focuses on the use of formulaic language by English as a second language students (L2). Research on the field of phraseology has shown that mastery of formulas is central for fluency and linguistic competence (Ellis, 1996). Studies on the use of formulaic language by native speakers (Ellis et al., 2008) have shown that native speakers process these structures as a single word. Considering the use of formulaic language by L2 students, research has shown that this can be problematic to learners as they do not know the correct word association (Men, 2018). This paper presents a literature review on the studies of formulaic language, more specifically of collocations, used by L2 learners. The first part of this paper deals with the different definitions of collocations, while the second part focuses on studies on collocation use by L2 learners.

KEYWORDS: collocations, formulaic language, corpus linguistics.

O uso de colocações em diferentes níveis de proficiência: uma revisão da literatura

RESUMO:

Esta revisão da literatura aborda o uso de sequências formulaicas por estudantes de inglês como segunda língua (L2). Pesquisas no campo de fraseologia têm mostrado que o domínio de sequências formulaicas é essencial para a fluência e a competência linguística dos alunos (Ellis, 1996). Estudos sobre o uso de sequências formulaicas por falantes de inglês como língua nativa (Ellis et al., 2008) mostram que falantes nativos processam estas estruturas como uma única palavra. Considerando o uso de fórmulas por alunos de L2, pesquisas mostram que estas podem ser problemáticas para os aprendizes, pois eles não sabem a associação correta de palavras (Men, 2018). Este artigo apresenta uma revisão da literatura em pesquisas sobre linguagem formulaica, mais especificamente em colocações, usadas por alunos de L2. A primeira parte deste trabalho aborda diferentes definições de colocações, enquanto a segunda parte apresenta pesquisa sobre colocações e seus usos por aprendizes de L2.

PALAVRAS-CHAVE: colocações, linguagem formulaica, linguística de corpus.

Corresponding Author:

LARISSA GOULART
<lg845@nau.edu>



Este artigo está licenciado sob forma de uma licença Creative Commons Atribuição 4.0 Internacional, que permite uso irrestrito, distribuição e reprodução em qualquer meio, desde que a publicação original seja corretamente citada. http://creativecommons.org/licenses/by/4.0/deed.pt_BR

1. INTRODUCTION

Several studies (Erman & Warren, 2000; Biber & Conrad, 1999; Pawley & Syder, 1983) suggested that language is mainly composed of fixed or semi-fixed language sequences. Aside from formulas pervasiveness in language, research on formulaic language processing showed that native speakers of English (L1) process these formulas as one single element. Ellis (1996, p. 111), for instance, argues that formulaic language is perceived as a “big word —the role of working memory in learning such structures is the same as for words”. Sinclair (1991, p. 110) agreed with this view, stating that formulas are a “single choice, even though they might appear to be analyzable into segments”. While Perkins (1999, p.56) explains the use of formulaic language saying that “the main reason for the prevalence of formulaicity in the adult language system appears to be the simple processing principle of economy of effort”, Wray (2005) argues that even though humans have the ability for processing language grammatically, or rather analytically, the preferred way of coping with language input and output is through chunks of language. In sum, previous research established that native speakers process formulaic language as a single word, nevertheless, these investigations do not account for how speakers of English as a second language (L2) use and process formulaic language.

Sinclair (1991) proposes that language users deal with formulaic language based on two principles, the open choice and the idiom principle. The first one allows for new and creative uses of language, while the second refers to the use of frequent combination of words. These two principles are especially important when considering speakers of English as an L2, as it is unclear whether they rely on the idiom principle or in the open choice principle when using their second language.

Ellis (1996) argues that L2 learners’ acquisition of formulaic sequences differs from that of native speakers, in the sense that native speakers process formulas relying on semantic associations, while L2 learners rely on orthography and phonology, driving them to, possibly, make wrong associations based on orthographic or phonological confusion. In a recent study, Ellis et al. (2008) confirmed that native speaker process formulas based on different criteria than L2 learners. While the latter used formulas that are more frequent, the former used formulas that had a stronger association between words.

In research about reading and writing, it is established that the use of formulaic language gives more fluency to a text (Ellis, 1996). Nevertheless, in L2 writing, different studies (Boers & Webb, 2018; Wray, 2013; Paquot & Granger, 2012; Nesselhauf, 2005) have shown that the use of formulaic language can be an issue for beginners as well as advanced learners, with proficiency level impacting the amount of formulas used as well as the types of formulas used. Another issue with the use of formulaic language in written texts has been raised by Yoon (2016), who argues that each register is characterized by the use of distinct formulas. Therefore, the aim of this literature review is to describe how language development influences the use of formulaic language in L2 students.

In order to address this goal, this paper is divided in five sections. Section two describes the approaches found in the literature for the study of formulaic language and discusses the different definitions used in phraseology studies. Section three presents the methodology used in this literature review. In section four the results of the literature review are presented in light of the research questions. Section five discusses the implications and limitations of this study.

2. DEFINING FORMULAIC LANGUAGE

In this section, the different approaches used on research about formulaic language are described along with the different terms used to refer to frequent strings of words.

2.1 APPROACHES TO THE STUDY OF FORMULAIC LANGUAGE

In the field of formulaic language, different terms have been used to define the same object of study, while sometimes the same term is adopted to define different objects of study. One of the reasons for this is the different approaches that have been used for the study of formulaic language. Wray (2005), Durrant and Mathews-Aydinli (2011), and Durrant (2014) describe three research approaches. The first one is the phraseological approach (e.g. Cowie, 1998; Ermand and Warren, 2000) which analyzes the meaning of a word combination. This approach is concerned with the degree to which the meaning of a word combination is predictable based on the meaning of its parts. It might also analyze whether words with similar meanings can be substituted in a phrase (e.g. *jump through hoops* or *skip* through hoops*). This approach usually relies on researchers' intuition of what is formulaic in a given language. Furthermore, Wray (2005) argues that this approach results in idioms rather than formulas. The second approach is the psychological approach (e.g. Wray, 2005; Ellis et al. 2008) which focuses on the mental processing and storage of language. This approach defines formulas as items, which speakers store and process as a whole. The third approach is the frequency one (e.g. Biber and Conrad, 2009; Hoey, 2005, etc) which focuses on the frequency of co-occurrence of certain linguistic combinations in a text. These linguistic combinations can refer to words, parts of speech, or semantic fields. The frequency approach is associated with corpus linguistics studies of formulaic language. Unlike the psychological approach, the object of study of the frequency approach are texts (written or spoken) produced by language users. One of the issues with this approach is that researchers have defined the limit of the string words being studied differently, producing different results.

Although these three approaches suggest that there are different phenomena being studied, Durrant and Mathews-Aydinli (2011) highlight the fact that the psychological and the frequency approach look at the same phenomenon from different perspectives. Wray (2005) also argues that the usage frequency of these formulaic sequences is associated to how they are stored and processed in the brain, thus corroborating Durrant and Mathews-Aydinli's (2011) argument. Furthermore, according to Henriksen (2013), nowadays many researchers adopt a combined approach using the frequency approach to find formulaic language, and then using their judgment to determine whether the words have a meaning relationship or using a frequency approach to determine the items to be tested based on a psychological approach. Therefore, both psychological and frequency approach will be taken into account in this literature review, while a phraseological approach, which focuses on idioms will not be addressed.

DEFINITIONS OF FORMULAS

One of the first definitions of formula can be found in Jespersen (1924/1976) who said that formulas "must always be something which to the actual speech instinct is a unit, which cannot be further analyzed or decomposed in the way a free combination can" (p.88). Later, Bolinger (1976) would say that "our language does not expect us to build everything starting with lumber, nails and blueprint, but provide us with an incredibly large number of prefabs" (p.1). Fillmore (1979) also said that "a very large portion of a person's ability to get along in a language consists in the mastery of formulaic utterances" (p.92).

From the late 80s onwards, a plethora of terms have been used to define formulaic language, many of these terms related to the development of corpus linguistics tools and new ways to analyze language. Wray (2005, p.9) presents all of the terms found to describe formulaic language in the figure below:

amalgams – automatic – chunks – clichés – co-ordinate constructions – collocations – complex lexemes – composites – conventionalized forms – F[ixed] E[xpressions] including I[dioms] – fixed expressions – formulaic language – formulaic speech – formulas/formulae – fossilized forms – frozen metaphors – frozen phrases – gambits – gestalt – holistic – holophrases – idiomatic – idioms – irregular – lexical simplex – lexical(ized) phrases – lexicalized sentence stems – listemes – multiword items/units – multiword lexical phenomena – noncompositional – noncomputational – nonproductive – nonpropositional – petrifications – phrasemes – praxons – preassembled speech – precoded conventionalized routines – prefabricated routines and patterns – ready-made expressions – ready-made utterances – recurring utterances – rote – routine formulae – schemata – semipreconstructed phrases that constitute single choices – sentence builders – set phrases – stable and familiar expressions with specialized subsenses – stereotyped phrases – stereotypes – stock utterances – synthetic – unanalyzed chunks of speech – unanalyzed multiword chunks – units

Figure 1 – Terms used to describe formulaicity (Wray, 2005, p.9)

While ideally a literature review concerned with the language development and the use of formulaic language would address all of the terms used, due to time and space constraints this literature review will focus only on collocations, which are defined below. This is because many studies on collocations and language development were published recently, while other definitions are not so prolific on this issue of language development. It is worth mentioning, though, that several recent studies (e.g. Staples et al, 2013; Huang, 2015) on lexical bundles have dealt with the issue of language development as well.

Durrant and Mathews-Aydınlı (2011, p.60) define collocations as

successions of linguistic entities that are best learned as integral wholes or independent entities, rather than by the process of placing together their component parts, either because (a) they may not be understood or appropriately produced without specific knowledge, or (b) because they occur with sufficient frequency that their independent learning will facilitate fluency

Men (2018) defines collocations as transparent in meaning (e.g. *make a decision*), rather than opaque as idioms (e.g. *raining cats and dogs*). While Durrant and Mathews-Aydınlı (2011) and Men's (2018) definitions deal mainly with meaning, it is worth taking into account the form collocations take. Collocations have restricted commutability, in other words the node word has a limited set of words that can co-occur with it (e.g. *commit a crime*). Furthermore, they are frequently strings of two or three words that occur with the following grammatical units: verb + noun, adjective + noun, preposition + noun, adjective + preposition, noun + noun, adverb + verb, adverb + adjective. It is not uncommon for researchers to deal with only one type of collocation, for example only verb + noun collocations, as will be discussed in this literature review.

Another important point regarding research on collocations is the spam of frequency. While some authors define collocations as two or three words that occur in adjacency, most definitions of collocations state that these words do not occur necessarily subsequently (Durrant and Schmitt, 2009). Words that co-occur can appear within a 4:4 spam, meaning that they can be separated by three words either to the right or to the left. Furthermore, collocates can be identified by association measures such as t-score or mutual information (MI). The use of these

two measures in order to define collocations was proposed by Durrant and Schmitt (2009). The t-score measures the certainty of an association between two words, emphasizing collocations that are very frequent. While the MI score indicates the degree to which two lexical items in a collocation occur more frequently than would be expected by chance.

For this literature review, collocations will be defined as any set of two or three words combined in which one of the words is a noun, adjective, adverb or verb, occurring adjacently or not. This definition is rather loose when compared to previous definitions, nevertheless one of the goals is to verify which methodology has been adopted in the research of collocations in L2 learners, therefore delimiting the definition would exclude papers that might be relevant.

Finally, considering the points raised in the discussion above, the guiding research questions for this literature review are:

RQ1 – How does proficiency influence the use of collocations by L2 student?

RQ2 – What methodologies have been used to investigate the use of collocations across proficiency levels?

RQ3 – In corpus linguistics research, which registers have been investigated?

The aim of this section was to present the different approaches used in the study of formulaic language, and to describe the definition used for this literature review. In the next section, the methodology used to gather the papers is described.

3. METHODOLOGY

The first step of the literature review was to search in a major database using the term: collocation* OR formulaic language AND language development. For this, I have used Web of Knowledge (using the *Arts and Humanities* filter). I have also used the same search (“collection*” or “formulaic language” +”language development”) on Google Scholar. With the results, I have analyzed the abstract of the papers in order to determine whether they presented research on the use of formulaic language by second language learners, and if so, if the papers mentioned the proficiency level of these learners. The second step was to go through the bibliography of the papers to check if there were any relevant studies to the research question. This literature review only used papers written in English.

The initial search resulted in 31 papers, nevertheless after further analysis some papers were excluded as they dealt with the effects of instruction in learners’ production of formulas, the use of idioms by L2 learners, or descriptions of collocation tools to be used by learners. Studies that were only based on interviews with students about their own perception of development (e.g. Barfield, 2008) were excluded. Furthermore, some studies did not report learners’ proficiency level, and were also excluded. Finally, 23 papers were analyzed for this literature review. Next section discusses the results of the literature review.

4. RESULTS

In this section the results of the studies on the use of collocations by L2 students will be discussed. It is worth highlighting that this overview takes into account specifically how these studies address the research questions, which means that other issues could be discussed based on the research reviewed on this papers.

In the total 23 papers were analyzed, most of them adopted a frequency approach to the study of collocations (Men, 2018; Paquot, 2018; Yoon, 2016; Crossley et al. 2015;

Bestgen & Granger, 2014; Granger & Bestgen, 2014; Namvar, 2012; Laufer & Waldman, 2011; Li & Schmitt, 2009; Durrant & Schmitt, 2009; Wang & Shaw, 2008; Nesselhauf, 2003; Altenberg & Granger, 2001; Granger, 1998). Surprisingly, from these papers only two (Durrant & Schmitt, 2009; Yoon, 2016) report the spam of words used in their search of collocations. Nevertheless, it is possible to assume from the examples of collocations provided in other studies that these researchers only took into account adjacent words (e.g. Paquot, 2018, Altenber & Granger, 2001).

As for the type of collocations studied, several papers investigated the use of verb plus noun (Men, 2018; Yoon, 2016; Namvar, 2012; Laufer & Waldman, 2011; Wang & Shaw, 2008; Wolter & Gyllstad, 2011; Bahns & Eldaw, 1993). Some studies focused on a variant of the verb plus noun collocation, the verb plus object-noun (Wolter & Gyllstad, 2011; Revier, 2008; Nesselhauf, 2003; Bonk, 2000). Another type of collocation studied was the adjective plus noun (Men, 2018; Granger & Bestgen, 2014; Namvar, 2012; Durrant & Schmitt, 2009; Koosha & Jafarpour, 2006) and the noun plus noun (Men, 2018; Granger & Bestgen, 2014; Durrant & Schmitt, 2009), which Durrant and Schmitt (2009) bundle together, calling it noun modifier collocation. Fewer studies looked into collocations with adverbs (Alzi'abi, 2017; Granger, 1998). Some studies conducted research on collocations involving prepositions (Namvar, 2012; Koosha & Jafarpour, 2006; Bonk, 2000). Unfortunately, a good number of studies did not report the type of collocation being investigated (Paquot, 2018; Crossley et al. 2015; Li & Schmitt, 2009; Koya, 2005; Martyńska, 2004.; Farghal & Obledat, 1995). Finally, Altenberg and Granger (2001) analyzed all words that collocated with the verb *make*, and other delexical verbs.

Even though it is not the focus of this paper, it is worth mentioning that some studies have compared the use of collocations between native speakers and L2 learners (Crossley et al, 2015; Granger & Bestgen, 2014; Laufer & Waldman, 2011; Wolter & Gyllstad, 2011; Durrant & Schmitt, 2009; Altenberg & Granger, 2001; Granger, 1998). These studies were included in this literature review because they clearly stated learners' proficiency level, and they showed the results for the L2 learners separately from the native speakers. Another issue that will not be addressed in the research questions but is present in some of the studies reviewed is a comparison between L1 backgrounds. Some studies separated their results according to learners' L1 (Crossley et al., 2015; Granger & Bestgen, 2014; Durrant & Schmitt, 2009; Wand & Shaw, 2008; Altenberg & Granger, 2001), while others only contained data from speakers of one L1 (Men, 2018; Paquot, 2018; Alzi'abi, 2017; Namvar, 2012; Laufer & Waldman, 2011; Wolter & Gyllstad, 2011; Li & Schmitt, 2009; Revier, 2008; Koosha & Jafarpour, 2006; Koya, 2005; Martyńska, 2004; Nesselhauf, 2003; Granger, 1998; Bahns and Eldaw, 1993).

In language development studies, most of them aimed at evaluating whether the use of collocations could be a predictor of proficiency level (Paquot, 2018, Crossley et al., 2015; Bestgen & Granger, 2014; Koosha & Jafarpur, 2006; Bonk, 2000). Table 1, below, presents a summary of all of the papers reviewed. The following subsections address each research question separately.

TABLE 1 – Studies on the use of collocations across proficiency levels

REFERENCE	SUMMARY OF GOALS	PARTICIPANTS	METHOD	RESULTS	REGISTER
Men (2018) - Vocabulary increase and collocation learning	This study focused on the developmental patterns of verb plus noun collocations across levels.	The Chinese Learner English Corpus was used. This corpus contains texts from learners across different levels of education, it totals a million words.	Corpus Linguistics: The author extracted all combinations of two words (verb plus noun, adjective plus noun and noun plus noun) in the corpus and compared its use to collocation dictionaries and the British National Corpus (BNC).	The results showed that while more advanced students produced more collocations, they also made more mistakes. It also showed that intermediate students use more delexical verbs plus noun collocations than advanced students.	Not mentioned
Paquot (2018) - Phraseological Competence: a missing component in university entrance language tests? Insights from a study of EFL learners' use of statistical collocations.	The aim of this paper was to show the relevance of collocations in the delineation of the Common European Framework of Reference for Languages (CEFR). The author used different measures of phraseological and lexical complexity, and collocational association to determine if these were good measures of learner proficiency.	The researcher used texts from the Varieties of English for Specific Purposes (VESPA) corpus, containing 98 research papers written by French students of linguistics. The papers were rated by three experienced teachers in order to determine proficiency level.	Corpus Linguistics: It is unclear how the author extracted the collocations.	The results indicated that, contrary to previous research, lexical and phraseological measures of complexity did not reflect proficiency level, while the correct use of collocations reflected learners proficiency level.	Research Papers
Alzi'abi (2017) - Guessing verb-adverb collocations: Arab EFL learners' use of electronic dictionaries	This research investigated how Arab EFL learners use verb plus adverb collocations based on two learner's dictionaries.	82 students of an MA in TESOL at a university in Jordania. The students were classified as advanced learners.	Dictionary task completion: Learners were given 22 verbs, 12 high-frequency and 10 low-frequency. Students were asked to provide adverb collocates for these verbs using only a learner's dictionary.	According to the authors, only 10% of the responses were appropriate, showing that even with the help of dictionaries learners were not familiar with the collocations for the verbs selected. Another finding was that subjects that used the dictionary had higher scores in the low-frequency verb collocates than those who did not use it.	
Yoon (2016) - Association strength of verb-noun combinations in experienced NS and less experienced NNS writing: longitudinal and cross-sectional findings	This study analyzed the use of verb plus noun collocations in a longitudinal corpus of inexperienced L2 writers in two registers. It also compared the use of collocations by L2 learners and native speakers.	The corpus of L2 learners was written by 51 high-intermediate students enrolled in an Intensive English Program. The corpus of native speakers was written by 46 MA TESOL students.	Corpus Linguistics: The author analyzed the use of verb plus noun collocations in a 4:4 spam in the corpus of L2 learners and of native speakers. The results were compared to those of the collocations found in COCA. The author used MI as an association measure.	L2 writers had different levels of collocation attainment across registers, performing better with narratives than argumentative essays. High-intermediate students also tended to use high-frequency collocations, usually with high frequency verbs (do, get, make, etc), while native speakers used three times more infrequent collocations than L2 students.	Narratives and Argumentative Essays
Crossley, Salsbury and McNamara (2015) - Assessing lexical profile using analytic ratings: a case for collocation accuracy	The aim of this study was to assess lexical proficiency in a corpus of oral and written texts produced by L2 students in three levels of English. One of the measures used for lexical proficiency was collocations.	The written corpus contains texts of 10 L2 learners in an Intensive English Program in the US. Eight L1s were represented. A matching L1 corpus was also compiled. The spoken corpus contained samples from 29 participants, in conversations between English L1 and English L2 students. A matching corpus of native speakers was compiled based on the Switchboard corpus.	Corpus Linguistics: The texts were rated by three native speakers of English who were graduate students in the English department. The authors defined collocations as an acceptable and expected sequence of words.	The results show that the use of collocation accurately is a predictor of raters scores in the written texts.	Free journal writing

REFERENCE	SUMMARY OF GOALS	PARTICIPANTS	METHOD	RESULTS	REGISTER
Bestgen and Granger (2014) - Quantifying the development of phraseological competence in L2 English writing: An automated approach	The aim of this paper was to establish if a different methodology to the extraction of collocations - CollGram - could be used to track the development of phraseological competence in L2 writing.	The authors used 171 essays written by 57 English learners at a university context. For the longitudinal analysis the authors compared the first and last essay written by each participant.	Corpus Linguistics: The author used the CollGram techniques, which is to analyze the use of bigrams in the corpus - regardless of parts of speech - and compare its association score based on the t-test and the MI score, with the same bigram in a large corpus, such as COCA or the BNC.	When comparing essays with higher grades and essays with lower grades, the association is stronger in the latter. The authors argue that the MI scores of the bigrams are positively correlated with the quality of the essays. Considering the longitudinal study, the analysis showed a significant evolution of the t-score between the first text and the last text written by students.	Argumentative Essays
Granger and Bestgen (2014) - The use of collocations by intermediate vs advanced non-native writers: a bigram-based study.	This research reported on a study of how intermediate versus advanced students used bigram collocations of different parts of speech.	The authors used 223 texts from the ICLE corpus. Learners language background were French (74), German (71), Spanish (78).	Corpus Linguistics: The researchers extracted bigrams of noun plus noun, adjective plus noun, and adverb plus adjective for the analysis. The authors used t-scores and MI to determine the association between the words.	The results indicate that the intermediate students use more high-frequency collocations and fewer lower-frequency collocations, even though these collocations are strongly related. Another major difference between levels is that advanced learners use more noun plus noun sequences than intermediate learners.	Argumentative Essays
Namvar (2012) - The relationship between language proficiency and use of collocation by Iranian EFL students	This study compared learners' knowledge of collocations based on writing scores and a on collocation test.	The students who participated in this study were 15 postgraduate students at a Malaysian university, with an intermediate level of English.	Corpus Linguistics and multiple choice task: This study investigates the use of adjective plus noun, noun plus preposition, verb plus noun and noun plus verb collocations.	The results showed that language proficiency, as assessed by the score obtained in the written essay, correlate to the use of collocations and learners score in the collocation test. Considering the different collocation types it seems that adjective plus noun, and noun plus verb collocations are more difficult for learners than verb plus noun and noun plus preposition collocations.	Personal Narrative (writing test)
Laufer and Waldman (2011) - Verb-Noun collocations in second language writing: a corpus analysis of learners' English	This paper reported on a research focusing the use of verb plus noun collocations in a corpus of L2 learners of English, whose first language was Hebrew.	The L2 students are represented by the Israeli Learner Corpus of Written English, which contains 759 argumentative essays written by school or university students. The corpus contains one essay per learner. Students level of proficiency was classified according to their level of study (e.g. 9th and 10th graders were considered basic, 11th and 12th intermediate and university students advanced). To represent native speakers the authors used a subset of the Louvain Corpus of Native English Essays (LOCNESS)	Corpus Linguistics: The authors created a list of high frequency nouns in the native speakers corpus and then extracted the collocations from these nouns using the corpus and collocational dictionaries. The same procedure was adopted for the L2 corpus. The focus is on verb plus noun collocations.	The results showed that native speakers produced almost twice as many collocations as the L2 learners. Taking into consideration a comparison across levels, the author noticed that advanced and intermediate levels produced significantly more erroneous collocations than beginners. The number of collocations produced also increased with proficiency.	Argumentative Essays
Wolter and Gyllstad (2011) - Collocational links in the L2 mental lexicon and the influence of L1 intralexical knowledge.	The aim of this paper was to analyze the perception of collocations by native speakers and L2 learners.	The participants of the research were 31 Swedish advanced learners of English and 37 native speakers who served as a control group.	Lexical Decision Task: All the collocations in this study were verb plus object-noun collocations. The authors of this paper created a list of 440 collocation items, with 99 items being real collocations and the other items being distractors or made up words and asked native speakers and L2 students to determine if the items were collocations or not.	The results showed that the response time of native speakers was shorter than that of L2 learners when determining whether a collocation was correct or erroneous. Furthermore, this study showed that L2 students were more successful with collocations that were congruent with their L1.	

REFERENCE	SUMMARY OF GOALS	PARTICIPANTS	METHOD	RESULTS	REGISTER
Li and Schmitt (2009) - The acquisition of lexical phrases in academic writing: a longitudinal case study.	In this research the authors followed one speaker of Chinese through her MA in TESOL at a British University. They have compiled a corpus consisting of all the papers she wrote for her classes.	One graduate students in TESOL, speaker of Chinese as an L1.	Corpus Linguistics and Interviews: A panel of three researchers analyzed by hand each collocation in the papers, establishing if they were appropriate or not. The authors also conducted interviews with the participant.	The results showed that her use of collocations increased by 37.4% from her first paper to her MA dissertation. The authors argued that she learned - or started to use 166 new collocations. Nevertheless, the amount of collocations used it is still less than that found on native speaker corpora.	Argumentative Essays and Critiques
Durrant and Schmitt (2009) - To what extent do native and non-native writers make use of collocations?	This research compared the use of collocations found in L2 learners text to that of native speakers. The goal of this paper was to test a new method to extract collocations from the corpus.	The researchers relied on three corpora. One containing texts written by graduate students in an EAP class in Britain, the other with texts written by undergraduate students in a Turkish university, and the third one was the Bulgarian subpart of ICLE. The native counterpart contained the same registers but written by native speakers, the authors also used LOCNESS for shorter essays like the ones found in ICLE.	Corpus Linguistics: The authors analyzed only premodifier-noun collocations, in other words adjectives plus noun or noun plus noun combinations. Only adjacent pairs were analyzed. Collocations were extracted based on basic concordance tools, but their strength was assessed based on their frequency in the British National Corpus (BNC) and based on the MI and t-score of the word combinations.	The results showed that lower-frequency collocations are more common in native speakers texts than texts from L2 students. Non-native speakers used more collocations with a stronger association.	Research Proposals and Argumentative Essays.
Wang and Shaw (2008) - Transfer and universality: collocation use in advanced Chinese and Swedish learner English	The purpose of this research was to describe the use of verb plus noun collocations in a corpus of advanced English learners.	The participants were 100 Chinese students and a 100 Swedish students at undergraduate level, who wrote short essays.	Corpus Linguistics: The authors chose the most frequent verbs in both corpora and analyzed their collocations in a collocation dictionary to determine if the collocations were correct or erroneous.	The lexical variety of the collocations was slightly higher for Swedish learners than for Chinese learners. However, the total occurrence of verb plus noun collocations was similar in both corpora, suggesting that regardless of L1, advanced students use the same amount of collocations. This study also showed that L2 learners of English use fewer collocations than native speakers.	Argumentative Essays
Revier (2008) - Evaluating a new test of Whole English collocations	The aim of this paper was to analyze the reliability and validity of a new test for learners' knowledge of collocation. Learners had their proficiency determined based on a background questionnaire and an English vocabulary test.	The participants were 56 Danish learners of English at different educational levels (1st and 2nd year of gymnasium and 1st year university students)	Multiple choice task: The test assessed learners knowledge of verb plus object-noun collocations.	The results showed that there was a significant difference between students at first level and third level, but the same did not occur with the second level.	
Koosha and Jafarpour (2006) Data-driven learning and teaching collocation of preposition: the case of Iranian EFL adult learners	This study had two main aims, the first one was to see if data driven learning would help students learn collocations with prepositions, the second one was to assess how proficiency played a role in the correct use of collocations with prepositions.	200 English major students at Iranian universities participated in this study. They were divided in different proficiency levels according to the Michigan. Students were divided in an experimental and a control group.	Translation task: Focusing on prepositions and their collocations (noun + prep, adjective + prep, prep + noun, verb + prep, prep + prep, and idiomatic expressions), especially those collocations that are difficult for Iranian learners.	The authors claim that collocation knowledge could be used as a measure of proficiency since the correct use of collocations correlate to learners' proficiency level.	
Koya (2005) - The acquisition of basic collocations by Japanese learners of English	This dissertation investigated how learners' passive and active knowledge of collocation related to level of proficiency.	The participants were 130 Japanese university students at four proficiency levels, which was measured through the vocabulary size test.	Multiple choice task and a translation task: The first one evaluated learners receptive vocabulary based on collocation dictionaries and the second one was a productive task of 68 collocations selected by the author based on previous research.	Learners with a bigger vocabulary used more collocations correctly. According to the author passive knowledge of collocations also correlated with productive knowledge of collocation.	

REFERENCE	SUMMARY OF GOALS	PARTICIPANTS	METHOD	RESULTS	REGISTER
Martyńska (2004) - Do English language learners know collocations?	The research presented in this paper identified the level of collocational competence of intermediate English learners, based on a test of collocation knowledge.	53 high school students with Intermediate English proficiency participated in this study.	Multiple choice task, cloze tasks, written translation task, and association task were used in this study.	The results showed that students were able to identify correctly 55% of the collocations. It also showed that years of L2 study had no impact in the collocation test.	
Nesselhauf (2003) - The use of collocations by advanced learners of English and some implications for teaching	This research paper reported on the results of an exploratory study, which analyzed the use of verb plus noun collocations in writing tasks.	The study was based on 32 essays written by German advanced learners of English for the International Corpus of Learner English (ICLE)	Corpus Linguistics: The author analyzed the use of verb plus object-noun collocations.	From all the verb plus object-noun combinations 846 were free combinations, 13 were idioms and 213 were collocations. From the collocations identified 56 were considered erroneous by native speakers.	Argumentative Essays
Altenberg and Granger (2001) - The grammatical and lexical patterns of make in native and non-native student writing.	This study investigated the use of collocations with high frequency verbs, especially make, by L2 students of English.	The researchers used the French part of the ICLE corpus, and essays written by Swedish students. Both students were classified as advanced. As a reference for native speakers use of collocations the corpus selected was the LOCNESS.	Corpus Linguistics: The authors analyzed the frequency of the verb make in the three corpora (Frel-CLE, Swedish students, and LOCNESS). The authors analyzed the collocations of make (in a span of 3 words to the right).	The results show that the use of make as a delexical verb proves to be difficult for advanced learners in both language backgrounds.	Argumentative Essays
Bonk (2000) - Testing ESL learners' knowledge of collocations	In this study the author sought to assess a collocation test, as well as determine whether language proficiency correlates to collocation proficiency. The researcher gave learners a reduced version of the TOEFL test and a version of the collocation proficiency test.	98 L2 students enrolled at University of Hawai'i participated in this study	Multiple Choice Task: The questions tested verb plus object-noun collocation, verb plus preposition and figurative-use-of-verb phrases.	The results showed a moderately high level of correlation between proficiency and collocational proficiency.	
Granger (1998) - Prefabricated patterns in advanced EFL writing: collocation and formulae.	The aim of this paper was to investigate the use of collocations involving adverbs in the ICLE corpus compared to the same structures in a corpus of native speakers.	This research used texts written by 56 French L1 students and 56 native speakers	Corpus Linguistics: The author created a list of the most frequent adverbs used with the meaning of amplification in both corpora and then analyzed the words that collocated with those adverbs in the corpora.	The results showed that learners used more collocations that were congruent with their first language. Nevertheless, in total native speakers had more occurrences of collocations with adverbs.	Argumentative Essays
Farghal and Oble-dat (1995) - Collocations: a neglected variable in EFL	The goal of this study was to show that basic collocations on the topics of food, color, and weather are a problem for learners of English. The authors created a list of 22 collocations in these topics and tested three levels of English users, asking them to complete the collocation.	The fill in the blank task was completed by 43 English majors. These were students in two levels, juniors and seniors. The translation task was completed by 23 English teachers.	Fill in the blank questionnaire and translation: The first test gave a word in English and asked the participants to produce its collocation, the second test gave an expression in Arabic and asked the participants to translate it to English.	The results showed that in both groups learners relied on synonyms, rather than using the appropriate collocation. Furthermore, for the translation task the advanced learners adopted different strategies, such as, paraphrasing and translating exactly as it is in their L1.	
Bahns and Eldaw (1993) - Should we teach EFL students collocations?	The purpose of this paper was to investigate the knowledge of verb plus noun collocations in advanced learners of English with German as an L1	The participants were 58 English majors at a university in Germany.	Cloze task and a translation task: These tasks focused on verb plus noun collocations	The results showed that in the translation task advanced students mistranslated 35.1% of the collocations. In the cloze sentence task more than 50% of the collections were erroneous.	

4.1 HOW DOES PROFICIENCY INFLUENCE THE USE OF COLLOCATIONS BY L2 STUDENT?

This section seeks to answer a question that seems rather simple: how proficiency influences the use of collocations by L2 students. However, the review shows that the answer to this question is not as simple as more proficiency, more collocations. All studies reviewed show that more advanced students do use more collocations - regardless of which part of speech the collocations investigated belong to. Nevertheless, there are more components to this question than previously anticipated. The first one is presented by Men (2018) and Revier (2008) which shows that it seems that there is a bigger increase in the use of collocations between lower level to intermediate level, than from intermediate to advanced. Another issue discussed in these papers is that even though advanced students use more collocations than lower level students, they do not show lexical variety (Wang & Shaw, 2008; Granger & Bestgen, 2014, Granger, 1998), this means that advanced students tend to repeat the same collocations. Yet, when compared to intermediate and beginner students, advanced students use more low-frequency collocations (Granger & Bestgen, 2014; Yoon, 2016) while these lower level learners use more collocations with high-frequency words. In addition, even though advanced students produce more collocations the amount of erroneous collocations found also increases with level (Men, 2018; Nesselhauf, 2003; Bahns & Eldaw, 1993). These authors noticed that L2 learners usually produce erroneous collocations within the same semantic field.

Although advanced learners are the ones that use more collocations, when compared to native speakers they produce a bit over half of the amount of collocations used by native speakers (Yoon, 2016; Laufer and Waldman, 2011). Therefore, these results sustain the claim that collocations are an issue for L2 learners even in more advanced levels.

In sum, the research described in these papers shows that the use of collocations increases with proficiency, nevertheless L2 speakers do not show collocation variety. Hence, when compared to native speakers, their use of collocation falls behind in a frequency count. Furthermore, another difference in the use of collocations between lower-level and higher-level learners is the use of low-frequency collocations, while higher-level learners tend to use more low-frequency collocations.

4.2 WHAT METHODOLOGIES HAVE BEEN USED TO INVESTIGATE COLLOCATIONS?

As discussed in section two, there are three approaches to the study of collocations. This review analyzed only two of them, the frequency and the psychological approach. The frequency approach, which is associated with corpus linguistics methods, was used in 14 out of the 23 papers reviewed. The remaining 9 papers adopted a psychological approach. In this section, the methods used for the extraction or analysis of the collocations are discussed.

From the papers that used corpus linguistics methods some tendencies can be observed, the first one refers to the extraction of all combinations of two or three words, without assigning a frequency threshold or any type of association measure. In several studies (Men, 2018; Crossley et al. 2015, Namvar, 2012; Laufer & Waldman, 2011; Wolter & Gyllstad; 2011; Li & Schmitt, 2009; Nesselhauf, 2003; Altenberg & Granger, 2001; Granger, 1998) it is unclear whether these researchers did not have any threshold measure or if they did not report it in their studies. This is especially important considering that if there was no frequency threshold then any combination of two words could be considered a collocation, if it appeared in the reference corpus, or was considered correct by the researcher. Considering association measures, only six studies (Men, 2018; Yoon, 2016; Bestgen & Granger, 2014; Granger & Bestgen, 2014; Durrant & Schmitt, 2009) report having used either MI or t-scores. The lack of frequency thresholds and association measures are an issue for studies of collocations as it makes the definition of this phenomenon quite broad and likely to impact the results of the studies.

Another point to be considered in collocation studies is the methodology used to determine whether a collocation was correct. Most of the studies that have compared the use

of collocations between native speakers and L2 students used corpora of native speakers in order to verify if the collocations were appropriate or not. Nevertheless, some studies relied on dictionaries (Men, 2018; Wang & Shaw, 2008) or a panel of judges (Li & Schmitt, 2009).

Considering that corpus linguistics is a methodology associated with frequency and computer tools, it seems that these researchers could adopt a centralized method of extraction of collocations and determining the appropriateness of each collocation.

Taking into account the psychological approach (Alzi'abi, 2017; Wolter & Gyllstad, 2011; Revier, 2008; Koosha & Jafarpour, 2006; Koya, 2005; Martyńska, 2004; Bonk, 2000; Farghal & Obledat, 1995; Bahns & Eldaw, 1993), most of these studies used translation tasks (Koosha & Jafarpour, 2006; Koya, 2005; Martynska, 2004; Farghal & Obledat, 1995; Bahns and Eldaw, 1993) from learners' L1 to L2 in order to determine if they knew the correct collocation with selected words. The second most used method was multiple choice tasks (Namvar, 2012; Revier, 2008; Koya, 2005; Martynska, 2004; Bonk, 2000), and few studies have used cloze tests (Azi'abi, 2017; Martynska, 2004; Farghal & Obledat, 1995; Bahns & Eldaw, 1993).

Summarizing the results of the literature review considering methodology, it seems that the problems with the definitions of formulaic language find their roots in the different methods used to extract and analyze collocations. Some of the issues are that translation tasks, for example, are dependent on learners L1; while papers on corpus linguistics research entails different methods of collocation extraction.

4.3 IN CORPUS RESEARCH, WHICH REGISTERS HAVE BEEN INVESTIGATED?

As briefly mentioned in the introduction, different registers use linguistic features, one of them being collocations, distinctively. Biber and Conrad (2009), for instance, explore register variation based on certain linguistic features extensively. The aim of this question is to determine which registers have been described based on the collocations used. This is relevant as the results regarding the use of collocations can be an indicative of the register being investigated, as well as of the learners' proficiency.

As we can observe in table one, argumentative essays is the most researched register in collocation studies (Yoon, 2016; Bestgen & Granger, 2014; Granger & Bestgen, 2014; Laufer & Waldman, 2011; Li & Schmitt, 2009; Durrant & Schmitt, 2009; Wang & Shaw, 2008; Nesselhauf, 2003; Altenberg & Granger, 2001; Granger, 1998) with one study focusing on other registers, such as, research papers (Paquot, 2018), narratives or narrative essays (Yoon, 2016; Namvar, 2012), free journal writing (Crossley et al, 2015), critiques (Li & Schmitt, 2009); and research proposals (Durrant & Schmitt, 2009). There are two issues with the low variety of registers being studied. The first one is that the findings reported in these studies might be associated with the register being investigated. As Yoon (2016)'s study shows, there is register variation in the use of collocations even between narrative and argumentative essays. The second one is that most of these argumentative essays being analyzed were written for English classes (Yoon, 2016; Crossley et al, 2015; Bestgen & Granger, 2014; Granger & Bestgen, 2014; Durrant & Schmitt, 2009), either English composition or programs in intensive English, which do not represent disciplinary writing. The studies that represent disciplinary writing (Paquot, 2018; Namvar, 2012; Li & Schmitt, 2009; Durrant & Schmitt, 2009; Koosha & Jafarpour, 2006; Nesselhauf, 2003; Altenberg & Granger, 2001) were written by English or Linguistic majors.

Furthermore, we can see an extensive use of the same corpus across studies, ICLE has been used in five studies (Granger & Bestgen, 2014; Durrant & Schmitt, 2009; Nesselhauf, 2003; Altenberg & Granger, 2001; Granger, 1998), with the French subpart being investigated in three studies (Granger & Bestgen, 2014; Altenberg & Granger, 2001; Granger, 1998).

In sum, even though there are plenty of studies on the use of collocations by L2 speakers,

this review shows that most of them investigate the same type of speakers, English students in their English classes; and the same registers, essays. This shows that there is a gap in research on collocations to investigate other registers, taking into account disciplinary writing.

The aim of this section was to present the results of the three research questions in light of the papers reviewed. The next section presents a discussion of the results and the limitations of this review.

5 DISCUSSION

After reading the studies reviewed, it is evident that Nesselhauf (2003) was right, when the author argued that collocational studies, even those using the same approach, have a hazy definition of collocations. This was evident by the different use of frequency measures, association measures, and even researchers' intuitions in the investigations reviewed. This literature review has mainly focused on language development studies, nevertheless, while reviewing this present paper, it became clear that there are other issues to be explored through a literature review. One of them being the role of L1 in collocation development. Some studies suggest that L1 plays a role in the correct use of collocations by learners (Men, 2018), while others suggest that regardless of L1, learners in the same proficiency level use collocations to the same extent (Wang and Shaw, 2008). A second issue that could be explored is how the parts of speech that form a collocation impact the acquisition of collocation; few studies (e.g. Men, 2018; Namvar, 2016) suggest that verb plus noun collocations are more difficult to be acquired than adjective or noun plus noun collocations.

Considering future research on collocations, it seems critical to report the word span in which the collocations analyzed occur, and also to adopt association measures in the extraction of collocations, otherwise any string of words occurring in the corpus will be considered a collocation. Furthermore, researchers could investigate other registers specially to describe register differences in the use of collocations, as the study conducted by Yoon (2016) did for narratives and argumentative essays.

Finally, the results of the literature review show that even for advanced learners collocations are an issue. With studies showing that, even though more proficient students use more collocations, they also make more mistakes when using these structures. This confirms that Nesselhauf (2002) is correct when suggesting that teachers should teach collocations explicitly, since learners usually see them as open choices, rather than units. That is, English teachers should not assume that learners will acquire collocations just by encountering them many times in written texts and class materials. These words that collocate together should be taught as units of language in English classes.

REFERENCES

- Altenberg, Bengt, & Granger, Sylviane. 2001. The grammatical and lexical patterning of MAKE in native and non-native student writing. *Applied linguistics*, 22(2), pg. 173-195. <https://doi.org/10.1093/applin/22.2.173>
- Alzi'abi, Safi Eldeen. 2017. Guessing Verb-Adverb Collocations: Arab EFL Learners' Use of Electronic Dictionaries. *Lexikos*, 27(1), pg. 50-77. <https://doi.org/10.5788/27-1-1394>
- Bestgen, Yves, & Granger, Sylviane. 2014. Quantifying the development of phraseological competence in L2 English writing: An automated approach. *Journal of Second Language Writing*, 26, pg. 28-41. <https://doi.org/10.1016/j.jslw.2014.09.004>
- Bahns, Jens, & Eldaw, Moira. 1993. Should we teach EFL students collocations?. *System*, 21(1), pg. 101-114. [https://doi.org/10.1016/0346-251x\(93\)90010-e](https://doi.org/10.1016/0346-251x(93)90010-e)

- Biber, Douglas, & Conrad, Susan. 1999. Lexical bundles in conversation and academic prose. *Language and Computers*, 26, pg. 181-190.
- Biber, Douglas, & Conrad, Susan. 2009. *Register, genre, and style*. Cambridge University Press.
- Boers, Frank, & Webb, Stuart. 2018. Teaching and learning collocation in adult second and foreign language learning. *Language Teaching*, 51(1), pg. 77-89. [https://doi.org/10.1016/0346-251x\(93\)90010-e](https://doi.org/10.1016/0346-251x(93)90010-e)
- Bolinger, Dwight. 1976. Again-one or two subjunctives?. *Hispania*, pg. 41-49. <https://doi.org/10.2307/339370>
- Bonk, William. 2000. *Testing ESL Learners' Knowledge of Collocations*. ERIC
- Cowie, Anthony. 1998. *Phraseology: Theory, analysis, and applications*. OUP Oxford.
- Crossley, Scott, Salsbury, Tom, & Mcnamara, Danielle. 2014. Assessing lexical proficiency using analytic ratings: A case for collocation accuracy. *Applied Linguistics*, 36(5), pg. 570-590. <https://doi.org/10.1093/applin/amt056>
- Durrant, Philip. 2014. Corpus frequency and second language learners' knowledge of collocations: A meta-analysis. *International Journal of Corpus Linguistics*, 19(4), pg. 443-477. <https://doi.org/10.1075/ijcl.19.4.01dur>
- Durrant, Philip, & Mathews-Aydinli, Julie. 2011. A function-first approach to identifying formulaic language in academic writing. *English for Specific Purposes*, 30(1), pg. 58-72. <https://doi.org/10.1016/j.esp.2010.05.002>
- Durrant, Philip, & Schmitt, Norbert. 2009. To what extent do native and non-native writers make use of collocations?. *IRAL-International Review of Applied Linguistics in Language Teaching*, 47(2), pg. 157-177. <https://doi.org/10.1515/iral.2009.007>
- Ellis, Nick 1996. Sequencing in SLA: phonological memory, chunking and points of order. *Studies in Second Language Acquisition* 18, pg. 91-126. <https://doi.org/10.1017/s0272263100014698>
- Ellis, Nick, Simpson-Vlach, Rita, & Maynard, Carson. 2008. Formulaic language in native and second language speakers: Psycholinguistics, corpus linguistics, and TESOL. *Tesol Quarterly*, 42(3), pg. 375-396. <https://doi.org/10.1002/j.1545-7249.2008.tb00137.x>
- Erman, Britt, & Warren, Beatrice 2000. The idiom principle and the open choice principle. *Text-Interdisciplinary Journal for the Study of Discourse*, 20(1), pg. 29-62. <https://doi.org/10.1515/text.1.2000.20.1.29>
- Farghal, Mohammed, & Obiedat, Houssein 1995. Collocations: A neglected variable in EFL. *IRAL-International Review of Applied Linguistics in Language Teaching*, 33(4), pg. 315-332. <https://doi.org/10.1515/iral.1995.33.4.315>
- Fillmore, Charles. 1979. On fluency. In *Individual differences in language ability and language behavior*, pg. 85-101. <https://doi.org/10.1016/b978-0-12-255950-1.50012-3>
- Granger, Sylviane 1998. Prefabricated patterns in advanced EFL writing: Collocations and formulae. *Phraseology: Theory, analysis, and applications*, 145, pg.160 - 175.
- Granger, Sylviane, & Bestgen, Yves 2014. The use of collocations by intermediate vs. advanced non-native writers: A bigram-based study. *International Review of Applied Linguistics in Language Teaching*, 52(3), pg. 229-252. <https://doi.org/10.1515/iral-2014-0011>
- Henriksen, Birgit 2013. Research on L2 learners' collocational competence and development—a progress report. C. Bardel, C. Lindqvist, & B. Laufer L, 2, pg. 29-56.
- Hoey, Michael 2005. *Lexical Priming: A new theory of language*. New York/ Abingdon, England: Routledge.
- Huang, Kaisheng 2015. More does not mean better: Frequency and accuracy analysis of lexical bundles in Chinese EFL learners' essay writing. *System*, 53, pg. 13-23. <https://doi.org/10.1016/j.system.2015.06.011>
- Jespersen, Otto 1976. *The philosophy of grammar*. New York: Norton.
- Laufer, Batia, & Waldman, Tina 2011. Verb-noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning*, 61(2), pg. 647-672. <https://doi.org/10.1111/j.1467-9922.2010.00621.x>
- Koosha, Mansour, & Jafarpour, Ali Akbar 2006. Data-driven learning and teaching collocation of prepositions: The case of Iranian EFL adult learners. *Asian EFL journal*, 8(4), pg. 192-209.

- Koya, Taeko 2005. *The acquisition of basic collocations by Japanese learners of English*. Unpublished doctoral dissertation, Wasedo University.
- Li, Jie, & Schmitt, Norbert 2009. The acquisition of lexical phrases in academic writing: A longitudinal case study. *Journal of Second Language Writing*, 18(2), pg. 85-102. <https://doi.org/10.1016/j.jslw.2009.02.001>
- Martyńska, Małgorzata 2004. Do English language learners know collocations?. *Investigationes linguisticae*, 11, pg. 1-12. <https://doi.org/10.14746/il.2004.11.4>
- Men, Hayan. 2018. *Vocabulary Increase and Collocation Learning*. Springer, Singapore.
- Namvar, Firooz 2012. The relationship between language proficiency and use of collocation by Iranian EFL students. *Language, Linguistics, Literature*, 18 (3), pg. 41 – 52.
- Nesselhauf, Nadja 2003. The use of collocations by advanced learners of English and some implications for teaching. *Applied linguistics*, 24 (2), pg. 223-242. <https://doi.org/10.1093/applin/24.2.223>
- Nesselhauf, Nadja 2005. *Collocations in a learner corpus*. John Benjamins Publishing.
- Paquot, Magali 2018. Phraseological Competence: A Missing Component in University Entrance Language Tests? Insights From a Study of EFL Learners' Use of Statistical Collocations. *Language Assessment Quarterly*, 15 (1), pg. 29-43. <https://doi.org/10.1080/15434303.2017.1405421>
- Paquot, Magali & Granger, Sylviane. 2012. Formulaic language in learner corpora. *Annual Review of Applied Linguistics*, 32, 130-149.
- Pawley, Andrew & Syder, Frances. 1983. Natural selection in syntax: Notes on adaptive variation and change in vernacular and literary grammar. *Journal of Pragmatics*, 7 (5), pg. 551-579. [https://doi.org/10.1016/0378-2166\(83\)90081-4](https://doi.org/10.1016/0378-2166(83)90081-4)
- Perkins, Michael (1999). Productivity and formulaicity in language development. In M. Garman, C. Letts, B. Richards, C. Schelletter & S. Edwards (eds), *Issues in normal & disordered child language: From phonology to narrative*. *Special Issue of The New Bulmershe Papers*, pg. 51-67.
- Revier, Robert Lee 2009. Evaluating a new test of whole English collocations. In *Researching collocations in another language*. Palgrave Macmillan, London, pg. 125-138. https://doi.org/10.1057/9780230245327_10
- Sinclair, J. 1991. *Corpus, concordance, collocation*. Oxford University Press.
- Staples, Shelley, Egbert, Jesse, Biber, Douglas, & McClair, Alyson 2013. Formulaic sequences and EAP writing development: Lexical bundles in the TOEFL iBT writing section. *Journal of English for academic purposes*, 12(3), pg. 214-225. <https://doi.org/10.1016/j.jeap.2013.05.002>
- Wang, Ying & Shaw, Philip 2008. Transfer and universality: Collocation use in advanced Chinese and Swedish learner English. *ICAME journal*, 32, pg. 201-232.
- Wolter, Brent & Gyllstad, Henrik 2011. Collocational links in the L2 mental lexicon and the influence of L1 intralexical knowledge. *Applied Linguistics*, 32 (4), pg. 430-449. <https://doi.org/10.1093/applin/amr011>
- Wray, Alison 2005. *Formulaic language and the lexicon*. Cambridge University Press.
- Wray, Alison 2013. Formulaic language. *Language Teaching*, 46 (3), pg. 316-334.
- Yoon, H. J. 2016. Association strength of verb-noun combinations in experienced NS and less experienced NNS writing: Longitudinal and cross-sectional findings. *Journal of Second Language Writing*, 34, 42-57. <https://doi.org/10.1016/j.jslw.2016.11.001>

Submetido: 09/05/2019

Aceito: 29/10/2019

Publicado: 31/12/2019

Autores:

Larissa Goulart

Doutoranda em Linguística Aplicada, Northern Arizona University (NAU).

E-mail: lg845@nau.edu

ORCID: <http://orcid.org/0000-0002-4190-4589>

Endereço: Department of English - Room 140 Building 18, Liberal Arts, 705 S. Beaver Street - Flagstaff, AZ, United States of America. ZIP CODE: 86011-6032.