A CORPUS-BASED INSTRUCTION TO LEARNING AND USING COLLOCATIONS IN A TRANSLATION AND INTERPRETATION PROGRAM: ANALYZING TRAINEE STUDENTS AND TEACHERS' PERCEPTIONS
UNA INSTRUCCIÓN BASADA EN CORPUS PARA EL APRENDIZAJE Y USO DE COLOCACIONES EN UN PROGRAMA DE TRADUCCIÓN E INTERPRETACIÓN: ANÁLISIS DE LAS PERCEPCIONES DE LOS ESTUDIANTES EN FORMACIÓN Y LOS PROFESORES

Rocío Rivera Cid
Universidad Andrés Bello, Chile
(r.riveracid@uandresbello.edu) (https://orcid.org/0000-0002-3776-3955)

	ABSTRACT
Keywords: collocations, collocational competence, corpus-based instruction, coca, translation and interpretation	This study examined the perceptions of college students from a Translation/Interpretation Program regarding the acquisition and application of collocations through the Corpus of Contemporary American English. Additionally, it explored the perceptions of teachers, translators and interpreters concerning collocational competence and a corpus-based instruction to using and translating collocations. A sample of 15 students participated in a corpus-based instruction to learn collocations through the corpus. Participants were required to complete a journal entry to reflect on their learning experience. In addition, they were asked to participate in a focus group once the intervention sessions were finished. Finally, two teachers, two translators, and one interpreter were interviewed to elicit their perceptions of collocational competence and a corpus-based instruction. The data collected were analyzed by means of Grounded Theory. Main findings revealed that learners recognized the value of learning collocations by means of a corpus-based instruction. Likewise, the teachers, translators and interpreters interviewed considered collocational competence and a corpus-based instruction to be significantly favorable. A more in-depth and extended study is advised to obtain more insight into the effectiveness of using corpora for the teaching and learning collocations in the field of Translation/Interpretation.
	RESUMEN
Palabras clave: colocaciones, competencia colocacional, instrucción basada en corpus, coca, traducción e interpretación.	Este estudio examinó las percepciones de los estudiantes universitarios de un programa de Traducción e Interpretación con respecto a la adquisición y aplicación de colocaciones a través del Corpus de Inglés Americano Contemporáneo. Además, exploró las percepciones de profesores, traductores e intérpretes sobre la competencia colocacional y la instrucción basada en corpus para el uso y traducción de colocaciones. Una muestra de 15 estudiantes participó en una instrucción basada en corpus para aprender colocaciones. Se pidió a los participantes que completaran una narrativa para reflexionar sobre su experiencia de aprendizaje. Además, se les solicitó participar en un grupo focal una vez

finalizadas las sesiones de intervención pedagógica. Finalmente, se entrevistó a dos profesores, dos traductores y un intérprete para obtener sus percepciones sobre la competencia colocacional y la instrucción basada en corpus. Los datos recopilados se analizaron mediante la Teoría Fundamentada. Los hallazgos principales revelaron que los estudiantes reconocieron el valor de aprender colocaciones a través de la instrucción basada en corpus. Asimismo, los profesores, traductores e intérpretes entrevistados consideraron que la competencia colocacional y la instrucción basada en corpus son significativamente favorables. Se recomienda realizar un estudio más profundo y extenso para obtener una mayor comprensión sobre la efectividad del uso de corpora en la enseñanza y el aprendizaje de colocaciones en el campo de la Traducción e Interpretación.

Introduction

Research on developing collocational competence in language learners and their mastery of formulaic language has gathered significant attention since the concept of collocation was first introduced (Nattinger & DeCarrico, 1992; Lewis, 1993). The claim is often made that the mastery of these formulaic expressions (collocations) is of paramount importance for communicative competence (Nation, 2001) and that it is fundamental to enable learners of a foreign language to process language fluently and idiomatically (Pawley & Snyder, 1983; as cited in Nation, 2001). Additionally, Ellis (2001) asserts that formulaic language plays a part in the mastery of language use proficiency. In the light of this, the benefits for learners of English as a second or foreign language of mastering such lexical items in relation to fluency support the need to conduct research focused on formulaic language such as collocations (Nation, 2001).

Nonetheless, this phenomenon is still not properly addressed by language teachers as an essential aspect to be taught along with grammar, phonetics, semantics, and syntax. This lack of interest is probably because of the dearth of teaching materials which are designed by using appropriate strategies and methodologies to develop the communicative competence, but which have a focus on grammar, phonetics, and other skills. Even though there is research which provides evidence on the importance of teaching collocations and developing learners' competence, there have been only a few studies aimed at investigating how learners of English as a second or foreign language develop this collocational competence. In this fashion, research suggests that there is a need to change the way English is taught in translation and interpretation courses from the current focus to an orientation towards vocabulary, especially collocations. In fact, there is research which supports a corpus-based approach to vocabulary and collocation teaching and learning due to the flourishing methodology of corpus linguistics which has been put to many different uses (Romer, 2009; Cobb, 1999).

The purpose of this research was to explore the perception of students from a Translation and Interpretation program regarding the learning and using of collocations using the Corpus of Contemporary American English (hereafter COCA). In addition, this study aimed to identify the perception of teachers, translators and interpreters of collocational competence and of a corpus-based instruction; the latter will aid in understanding their stance on teaching and learning collocations, especially within the field of translation. The aims of this study relate to the belief that translators and interpreters need to develop collocational competence for accurate translation, and that a corpus-based approach to learning collocations and translating them is an effective one for translators.

This investigation followed three main research questions, namely (a) What are the perceptions of students from a Translation/Interpretation program regarding the learning and using of collocations? (b) Do students from a Translation/interpretation program perceive a corpus-based instruction as appropriate for the learning and using of collocations? and (c) What are teachers and professional translators/interpreters' perceptions of collocational competence and a corpus-based instruction to learning and using collocations? In addition, the general objective of this investigation was to explore the perceptions of students from a Translation and Interpretation program regarding the learning and using of collocations using the Corpus of Contemporary American English, and well as to explore the perceptions of teachers, translators and interpreters of collocational competences and a corpus-based instruction to using and translating

collocations. Finally, the specific objectives for this investigation were (a) To identify students' perceptions regarding the learning and using of collocations; (b) To determine the extent to which students perceive a corpus-based instruction as appropriate and useful; and (c) To identify teachers, interpreters and translators' perceptions of collocational competence and a corpus-based instruction to learning and using collocations.

This article first presents the state-of-the art through a literature review. Then, the research methodology and procedures, as well as the analysis of the data, are described. Finally, this article presents the analysis of results, discussion, and main conclusions.

The concept of collocational competence has received much attention since Firth (1957) first coined the term collocation. Since then, the literature has shown that teachers and materials designers have focused on the need to develop a collocational competence and there is awareness that this language component should be addressed explicitly in language instruction (Howarth, 1998a; 1998b). Although there is vast literature on the Lexical Approach and on the importance of learning collocations rather than memorizing individual words, it is still difficult to find a precise definition of the term collocation. Most definitions seem to converge on the co-occurrence of words. Firth (1957), for example, claims that "you shall know a word by the company it keeps." (p. 179). In the same line, Nation (2001) believes that "knowing a word involves knowing what words it typically collocates with." (p. 74). Collocations can then be defined by Hill (2000) as the "the way words combine in predictable ways." (p. 48). The author asserts that, although teachers are familiar with the concept of communicative competence, it is necessary to extend this to collocational competence in the sense that it is a requirement for learners to master the target language in oral and written production. In addition, the author believes that the lack of competence in this area leads learners to make mistakes since they seem to be forced to create long utterances to compensate for the lack of collocational expressions to mean precisely what they want to say.

The importance of collocational competence is incontrovertible. As it is acknowledged by Fan (2009), "the importance of collocational competence is beyond dispute. It enables the learners to speak more fluently, makes their speech more understandable and helps them write or sound more native-like." (p. 111). Nattinger and DeCarrico (1992) support this and add that formulaic expressions, including collocations, are at the heart of language acquisition and thus help students improve speech, listening, reading and writing. Pawley and Syder (1983; as cited in Nation, 2001) also consider that the best explanation as to how language users can choose appropriate ways to say things from a wide range of options (native-like selection) and can use the language fluently (native-like fluency) is that some units of language are stored as chunks in memory.

To this regard, research has shown that language learners need to have a working knowledge of collocations to produce fluent and native-like discourse (Ellis, 1996; Conklin & Schmitt, 2008; Howarth, 1998a; 1998b). Similarly, Conklin and Schmitt (2008) agree that word combinations form a large part of any discourse; consequently, they need to be part of the vocabulary repertoire of anyone who attempts to learn a second or foreign language. In addition, Lewis (1993) suggests that native speakers of a language have a large repertoire of lexical chunks which are of key importance for the fluent production of the target language. Thus, fluency does not depend on grammar structure rules and a set of isolated words, but on the efficient access to this stock of lexical units which facilitate language production (Lewis, 1993; 1997a; 1997b; Thornbury, 2002). As Wilkins (1972; as cited in Thornbury, 2002) points out, "without grammar very little can be conveyed, without vocabulary nothing can be conveyed" (p. 13). This is also

acknowledged by Richards and Rodgers (2001) who believe that "the building blocks of language learning and communication are not grammar, function notions, or some other unit of planning and teaching but lexis, that is, word and word combinations" (p. 132). Therefore, it is on these language unit of language that teachers should center, though it is not unusual to see language teachers who fix their attention to paradigmatic relationships (lexical items which can be substituted for other ones in lexical and grammar contexts—synonyms) rather than syntagmatic relationships or the co-occurrence of lexical and grammatical items, known as collocations.

Even though it is a well-known fact that EFL students face problems in achieving native-like language due to a lack of collocational awareness, few studies address this issue empirically and the ones that do mainly focus on the use of these formulaic expressions in receptive rather than productive language (Howarth, 1998a; 1998b). In addition, there is scarce research which has shed light on language learners' collocational competence in terms of the way they process collocations and what teaching strategies are used in the classroom to help students understand and use them in oral and written discourse. What has been found, in this respect, is that some grammatical errors or mistakes are believed not to disrupt communication in the target language, whereas lexical errors are said to interfere with meaning, resulting in a lack of understanding between participants of an interaction (Gass & Selinker, 2001).

Conklin and Schmitt (2008), for example, investigated the advantage of formulaic sequences by comparing the reading times of these sequences and the matched non-formulaic phrases for native and non-native speakers. They found that formulaic phrases were read more quickly than non-formulaic sequences, concluding that "formulaic sequences have a processing advantage over creatively generated language." (p. 72). These results provide evidence to support the benefits of a collocation-based syllabus for EFL learners.

In the same fashion, Howarth (1998a) intended to identify and analyze non-standard phraseology (lexical collocations) in non-native academic writing. The findings suggest that "native speakers employ about 50 per cent more restricted collocations and idioms (of a particular structural pattern) than learners do, on average" (p. 177) and that this might "reflect learners' general lack of awareness of the phenomenon" (p. 186). This is relevant for the present study in that it provides evidence of the lack of collocational awareness of EFL students, which might affect their language production in terms of fluency and proficiency.

Other investigations relate to the study of the use of collocations by ESL students following a task-based approach. Fan (2009) attempted to understand the use of collocations of ESL secondary school students in Hong Kong. Some of the findings show that there is a need for an in-depth understanding of collocational competence and an innovative pedagogical approach to the learning and teaching of collocations. If this is the case, once again there is evidence that collocations play an important role in the language performance of a non-native speaker.

Several studies show the benefits of using a DDL approach to language learning (Cobb, 1999, 2003; Horst, Cobb, & Nicolae, 2005). However, only few studies have attempted to investigate the relation between learning and using collocations through a corpus-based approach and the effectiveness of corpus-based strategies and techniques to teach vocabulary (Cobb, 1999)—the majority of studies aim at describing the benefits of concordance for language learning (see for example Johns, 1991). This is why it was necessary to present the students of this study with a different approach to collocation learning and using, namely, a corpus-based approach.

The corpus-based approach in the field of language education and Linguistics has gained prominence since the mid-1980s. In this regard, much has been written about the positive influence of corpus linguistics on teaching foreign languages (Cob, 1999, 2003; Johns, 1991; Bernardini, 2004; Romer, 2009; Granath, 2009). There is no doubt that the corpus-based approach has been proven to be ground-breaking in teaching/learning languages (Bernardini, 2004). This is because the impact of bringing corpora into the classroom has been related to a "shift from deductive to inductive learning routines" (Bernardini, 2004: 16), which means students are encouraged to discover the language rather than learn it by memorizing grammar/lexical patterns.

Recent studies have further emphasized the importance and effectiveness of corpora tools in developing lexical knowledge and competence. For instance, a study conducted by Mohammed (2022) has demonstrated how translation trainees can benefit from monolingual, comparable and parallel corpora to improve fluency, accuracy and instrumental competence. This study highlights the integration of corpora tools into translation training and demonstrates thepositive perceptions of such tools to enhance professional readiness.

Furthermore, Romer (2009) claims that "corpus linguistics can make a difference for language learning and teaching and that it has an immense potential to improve pedagogical practice" (p. 84). Despite the many benefits of the use of corpora to enhance language learning, "the regular use of corpora in the EFL classroom is still a rare occurrence." (Granath, 2009: 47).

In addition to the corpus-based approach, a significant amount of literature is also available concerning DDL and its status in language teaching and learning. To this regard, Johns (1991) has asserted that language learning implies "the notion that the task of the learner is to discover the foreign language and that the task of the language teacher is to provide a context in which the learner can develop strategies for discovery [...]" (as cited in Johns & King, 1991: 1). Data Driven Learning (DDL) is defined by Johns (1991) as an approach in which the language student is driven by the access to linguistic data in their learning process.

Since corpus-based activities began gaining some advocacy on behalf of language teachers, there has been a growing number of publications aiming at presenting a range of uses for this approach in the classroom for vocabulary teaching/learning. In fact, many researchers have highlighted the benefits of using DDL to teach vocabulary to second or foreign language learners (Cobb, 1999, 2003; Horst, Cobb & Nicolae, 2005). These studies have demonstrated that some traditional activities such as gap filling can be improved if they are based on authentic texts. Furthermore, besides improving teaching materials, research has shown that direct DDL activities where learners use corpora by themselves have proven to be effective for vocabulary learning (Cobb, 1999). To this regard, Cobb's work (1997, 1999) is one of the few studies which has empirically proven the effectiveness of corpus-based strategies and techniques to teach vocabulary (Cobb, 1999, 2003). In his study, Cobb (1997) was able to evaluate the vocabulary learning outcomes from his students when they learned new words using concordance. The results found that students using the concordancer experienced limited but steady vocabulary Growth. In a follow up study, Cobb (1999) found that concordance also facilitates the acquisition of transferable word knowledge since the subjects in the study were able to use their knowledge of words in different activities and a range of other contexts.

Recent research has shown the impact of DDL on enhancing learners' language proficiency, particularly in the acquisition of noun collocations. He and Xie (2024) explored how DDL promores learners' autonony in language learning. By engaging with

authentic data from corpora, learners are able to independently discover collocational patterns and apply them effectively in written and spoken contexts. Their study showed that DDL not only improved collocational competence but also fostered learner motivation and self-directed learning. Participants in the study demonstrated improved proficiency in producing natural-sounding collocations, a skill which proves to be vital for Translation and Interpretation students. This is in line with the growing body of research advocating for the use of corpora to enhance collocational awareness in specialized fields where accuracy in language use is essential. As such, DDL is a solution for improving collocational competence, as it offrs linguistic and cognitive benefits for learners to master the target language.

While the effectiveness of corpus-based approaches in for vocabulary and language acquisition had been well-documented (Cobb, 1999; Bernardini, 2004), recent studies have also explored the role of blended corpus-based instruction in improving writing proficiency. For instance, Satchayad and Charubusp (2022) demonstrated that such instruction significantly enhances the writing proficiency of Thai university students by using corpus tools within interactive learning environments. Their study proves the potential of combining authentic corpus data with classroom activities to promote better language proficiency outcomes.

Unfortunately, time constraints are also to be considered when using corpora since it is, in most cases, difficult to find the time to integrate corpus research and exploration into language courses (Granath, 2009). This is said to be the main stumbling block for using and integrating corpora in the language classroom. As Granath (2009) explains, "the main problem with incorporating training in using corpora into regular courses is, as far as I can see, that only a limited amount of time can be set aside for actual hands-on exercises in the language lab" (p. 55).

The Corpus of Contemporary American English, created by Mark Davies from Brigham University in 2008, is the largest genre-balanced corpus of any language which has been designed to track and study changes in language (Davies, 2010). This corpus is made of more than 450 million words of text which are divided among spoken, fiction, popular magazines, newspapers, and academic texts. It includes 20 million words each year from 1990 to 2012, and it is a corpus suitable for looking at ongoing changes in language (Davies, 2010). The most important characteristic of this corpus is that it is considered to be a "monitor corpus", opposite to a static one which is not updated once it has been created, such as the British National Corpus. Hence, the COCA is a dynamic collection of texts to which new ones are added. Davies (2010) explains that the goal of this corpus is to "allow users to search the continually expanding corpus to see how language is changing." (p. 447). Although this has been the goal of many corpora, the Corpus of Contemporary American English is the only reliable monitor corpus used to carry out research of linguistic change in contemporary English.

Even though Corpus Linguistics and Data-Driven Learning have gained prominence in the last decades, there is still a gap concerning to what extent the COCA is effective for vocabulary/collocation teaching and learning. One attempt to bridge this gap is the work by Shaw (2011) which intends to provide teachers and students with a resource book where they can access information regarding the use of COCA for different vocabulary activities. In her work, Shaw (2011) presents different tasks for vocabulary learning by using the features available in the interface, such as word frequency activities, parts of speech, morphological knowledge, synonym knowledge, and collocational knowledge. In her book, Shaw (2001) suggests collocation searches are one of the most useful features of the corpus since they engage students in DDL activities, among other

advantages. This resource book presents many applications of the corpus (COCA) for vocabulary teaching/learning; however, it is by no means an empirical investigation on the effectiveness of this corpus on vocabulary/collocation learning. This is why the present study contributes to bridging the existing gap in research concerning a corpus-based approach (COCA) to collocation learning.

Research Design

This study was conducted using a qualitative paradigm, following a descriptive and exploratory design. This paradigm corresponds with the data-collection instruments selected for the data elicitation in this study, namely, interviews, a focus group and reflective journals. The reason for this was to identify students' perceptions of learning and using collocations by the Corpus of Contemporary American English, as well as the perceptions of some teachers and professionals regarding collocational competence and the use of corpora in a Translation and Interpretation Program at university level.

A purposive sample of 15 students from the Translation and Interpretation program at Pontificia Universidad Católica de Valparaíso (PUCV) participated in this research. As part of their academic requirements, the students involved in this study aim at reaching C1 level as stipulated in the Common European Framework of Reference for Languages (CEFR) during their second academic semester of their second year. The CEFR consists of a set of guidelines used to describe learners' achievements in foreign languages across Europe and in other countries.

For this research, 3 different data collection instruments were used: journal entries used during the pedagogical intervention sessions, a focus group and interviews conducted after the intervention. Due to the need for eliciting learners' internal processes and thoughts while participating in the pedagogical intervention sessions, reflective journals were used. These journals are often used to have learners write about their language learning experience without the pressure of framed questions designed by the researchers (Mackey & Gass, 2005). The journals were required from students to fulfill the first and second specific objectives of this study, namely, to identify learners' perceptions of the learning and using of collocations and their perception on the appropriateness of a corpus-based instruction. This retrospective assessment tool involved a journal entry which students had to complete in Spanish after each lesson during the pedagogical intervention to reflect on their learning process while participating in a corpus-based instruction. Students were asked to write their entries in Spanish to avoid focusing on language rather than on the learning process.

To fulfill the first and second specific objectives of this study, a second instrument, a focus group session, was also conducted. For this session, a group of students was randomly chosen to participate in the discussion. The questions for this focus group emerged from the need to gather in-depth information about learners' perceptions of the learning and using of collocations and on the use of COCA for collocation learning.

To fulfill the third specific objective of this study—to identify teachers, interpreters and translators' perceptions of collocational competence and a corpus-based instruction to learning and using collocations—two teachers of English, two translators and one interpreter were interviewed by means of semi-structured interviews designed for this study. The semi-structured interviews were piloted before they were administered, for validation purposes.

To collect the necessary data, a series of steps were carried out. Firstly, students were invited to an induction session where the objectives of this study were explained. Secondly, a consent form adapted from Mackey and Gass (2005: 323) was given to the students to comply with the ethical issues required in any study. By means of a short presentation and the information presented in the consent form, the students were informed of the objectives of this study, as well as the confidentiality of the data. Then, a pedagogical intervention was designed with the objective of introducing learners with a corpus-based approach to learning and using collocations. In each of these sessions, learners were instructed on the use of COCA to learn collocations and to use them in written tasks. After each intervention session, learners were asked to register in their journal entries their thoughts on the use of the corpus and their learning experience through this type of instruction. Following this, a group of students participated in a focus group with the objective of gathering in-depth information about their perceptions of learning and using collocations through COCA. Finally, three semi-structured interviews were designed. An interview was held with an English teacher currently teaching the course, aiming to gather information about her views on collocational competence and the use of corpora by translation students at PUCV. A second interview was conducted to a teacher of English who teaches academic writing with the objective of gathering information regarding his stance on collocational competence and a corpus-based instruction. A third and fourth interviews were conducted with two professional translators with the aim of gathering information about the importance of the development of collocational competence and the use of corpora for translation. One of the professional translators was also an expert in terminology who oversees the courses Terminology 1 and 2 at the same program; the other professional translator is a former student from PUCV. The fifth interview was conducted with a professional Interpreter with the objective of eliciting information about the importance of the development of collocational competence and the use of corpora in the professional field.

As the research conducted was qualitative, there was no statistical analysis considered for the interpretation of the data gathered for this study; rather, the researcher looked for students and professionals' perceptions of learning and using collocations using the COCA. The data obtained from the focus group, interviews and journal entries were transcribed and analyzed using Grounded Theory through an induction process. This means that the data gathered is coded through a constant comparative method which is used for analyzing data to develop theory. This method was appropriate for this study since it can be applied to data units of every size (Glaser & Strauss, 1967). The data obtained from each instrument revealed categories which will be presented in the analysis and discussion section. The categories analyzed emerged from the codification of the data-collection instruments which were separately coded. This means that the categories were not pre-conceived, but rather revealed in the process of data coding. For the sake of this analysis, only open and axial coding as proposed in Grounded Theory was conducted due to the small-scale nature of this study.

Analysis and discussion of results

Students' perceptions regarding the learning and using of collocations
From the students' reflections, one clear category during the codification process was identified during the codification process: 'importance of learning and using collocations'. In this category, it was possible to elicit students' perceptions regarding two

298

aspects of the learning and using of collocations. The first aspect analyzed was that of students' awareness of collocations. One of the fundamental principles of the Lexical Approach relates to the existence of prefabricated language units which account for a significant portion of a native speaker's discourse. Hence, students' "awareness of their ability to chunk language successfully is key to mastering the target language" (Lewis, 1993: vi). To this regard, students' reflections show that collocational awareness is something they achieve at university, rather than at school. From the interviews and journals, it was possible to elicit learners' perceptions in relation to the difficulty to understand what collocations are, probably because of their lack of awareness of collocations when they took their first language course at university. To this, Hill (2000) explains that "It is likely to be helpful to make learners explicitly aware of the lexical nature of language" (p. 161).

It was also possible to observe that collocational competence as a concept is then something learners are generally not aware of unless they are part of a teacher-training program or a translation and interpretation course. Although this is still an assumption, it is possible to affirm that students do not fully understand the importance of this competence until they are explicitly taught what this is and, consequently, they risk making language mistakes because of their lack of collocational competence. To this, Hill (2000) asserts that "the reason so many students are not making any perceived progress is simply because they have not been trained to notice which words go with which" (p. 14).

The second aspect which emerged from this category is that of the importance of learning and using collocations. Students recognized several advantages of learning and using collocation. The first advantage relates to *using lexical chunks to achieve language proficiency*. Lewis (1993) explains that lexis (including collocations) is central to language proficiency since the learning of lexical items fulfills students' need to go beyond the "intermediate plateau"—a prolonged stage of interlanguage development in which learners appear to be stuck due to their lack of collocational awareness and knowledge. The second advantage of learning and using collocations relates to *reading comprehension skills*. Receptive skills have been proven to be influenced, either positively or negatively, by the knowledge of collocations (Nattinger & Decarrico, 1992). This is because comprehension is either hindered or enhanced by the knowledge that the reader has in terms of using collocational expressions. The third advantage of using and learning collocations has to do with *the quality of translation*. That is, students acknowledge the importance of collocations since they believe translators/interpreters need to use collocations in oral and written English to achieve good-quality translations/interpretations. Indeed, collocations are of great relevance for the quality of a translation (Manafi & Ghaffarof, 2013).

The fourth advantage recognized by students was related to *using and learning collocations to improve fluency in the target language and to achieve near native-like English*. Students found collocations helpful, as they made speaking in the target language more fluent and also made their English sound more like that of a native speaker. To this regard, Fan (2009) explains that the importance of collocational competence is essential to achieve fluency and to native-like English. Nattinger and DeCarrico (1992) also acknowledge the importance of collocations for fluency since this aspect of the language is identified as one of the advantages of learning and using collocations in the target language. It is then clear that collocations are a tangible marker of non-nativeness, thus, students produce a more native-like discourse if they use collocations in their oral and written language production (ibid).

Students' perceptions of a corpus-based instruction for learning and using collocations

There is plenty of research regarding the influence of corpus linguistics on teaching foreign languages. There is also vast literature available concerning DDL and its status in language teaching and learning. However, there isn't much investigation concerning the effectiveness of COCA in vocabulary/collocation learning. The data gathered for this study show that these learners' perceptions of the use of COCA is highly positive since they recognized many of its benefits, namely, using COCA (1) for the learning of collocations, (2) for active learning, (3) for learning domain-specific terminology, (4) for written production improvement and (5) for translation/interpretation.

The first benefit recognized by students is that the COCA is greatly valued for collocation learning since the corpus aids students' awareness on the importance of collocations. This is supported by Shaw (2011) in that she recognizes that one of the benefits of using COCA is that teachers can create activities for collocational awareness and learning by using different features of the online corpus. In this sense, she agrees that collocation searches are the most useful features of the corpus since they engage learners in corpus-based activities.

A second benefit of using COCA is that this corpus encourages active learning, thus it encourages students to learn vocabulary, especially word combinations, in a different way, other than memorizing vocabulary guideline**s**. To this regard, it has been proven that the more active a learner is in their learning process with regards to vocabulary, the easier it is to recall the lexical items and combinations when needed (Johns, 1991; in Johns & King, 1991). This is also supported by Johns (1991) since it is his firm conviction that learners should "learn how to learn" (as cited in Johns & King, 1991: 1).

In addition to learning a wide range of general collocations, learners who participated in this study recognized a third benefit of using COCA, namely, learning about collocations as domain-specific discourse. This is recognized as a benefit by students since through the corpus they can have access to certain characteristics of words and expressions (such as frequency of occurrence) which make them likely to be used in certain text types and genres. To this, Nation (2001) points out that "it is thus possible to have special vocabularies for speaking, for reading academic texts, for reading newspapers, for reading children's stories, or for letter writing." (p. 17). In this regard, it is extremely useful for learners to use corpora since not only are they attractive as a learning tool, but they are also useful for specialized terminology which is part of learners' competences to acquire as future translators and interpreters.

Domain-specific collocations also depend on their frequency of occurrence in various genres. To this respect, the COCA is highly valuable since it provides learners with frequency-related data for collocations. Hence, learners are not only able to identify patterns of lexical use, but also the different contexts in which some combinations of words are used in diverse genres and registers.

A fourth benefit perceived by the students was the use of COCA for written production improvement. It is a widespread belief that an advanced user of a language must be able to produce a written discourse which is rich not only in complex grammar structures, but also in vocabulary, especially in collocations (Nattinger & DeCarrico, 1992). In fact, Hill (2000) recognizes the importance of using multi-word units in that complex ideas are often expressed lexically, and not grammatically. In their reflection on the use of COCA, learners manifested the importance of using this tool since it is of great help to develop written production.

Finally, a fifth benefit recognized by learners was the use of the COCA for translation and interpretation (English to Spanish and inverse translation/interpretation). This was highly valued by the students since translation is not always straightforward in terms of the meaning of lexical and grammatical collocations which do not have an equal term in Spanish. In this sense, it was useful for learners to search for word combinations in the corpus by using the feature KWIC (key word in context). This way it was easy to understand the meaning of the collocation in context to translate it. It is relevant to mention that no research has been conducted regarding the use of the COCA for translation/interpretation; rather, all studies which relate corpora and translation aim at investigating the use of comparative, multilingual and parallel corpora (Baker, 1992).

Not only is the corpus useful for translating from English into Spanish, but it is also useful for inverse translation. As far as this type of translation is concerned, learners agreed that the use of the corpus is highly favorable since it is easier to translate a concept into Spanish by, for example, comparing a collocation in terms of the different registers in which it might be frequently used.

Along with the benefits, the participants in this investigation recognized drawbacks in the use of the corpus. The participants in this study agreed that the use of the COCA is, in general, time-consuming. In this respect language teaching experts do not seem to agree since they see the use of corpora as an efficient tool in terms of speed (Cobb, 2003). What was evidenced in this study is that all participants considered it a difficult tool to use in class since they were not familiar with it; however, once they had time to work with it, it was much easier to use. Despite this, the learners still thought the corpus took much time to use.

Teachers, translators and interpreters' perceptions of collocational competence and a corpus-based instruction to learning and using collocations

A teacher of English who was an expert in EFL writing, a teacher of general English, a professional interpreter, and two professional translators were asked about their stance on the importance of collocational competence for translators and interpreters. Their perceptions were found to be related to two different aspects, namely, (a) the importance of teaching collocations and developing collocational competence and (b) the importance of collocational competence for translators and interpreters in general and domain-specific translation/interpretation.

Regarding the first aspect—the importance of teaching collocations and developing collocational competence—both teachers and professionals in the translation and interpretation field agreed that collocations are one aspect of language learning which has long been neglected in terms of its importance both at school and university levels. Moreover, they agreed on the importance of teaching collocations and developing collocational competence in language learners. It was also possible to evidence the significance of vocabulary from the perspective of a teacher whose job is to teach strategies for students to become proficient in the written production of the target language. It is the teacher's belief that vocabulary is sometimes overlooked by some fellow colleagues who prioritize grammar and not lexis. This is supported by Ellis (1997) in that he states that language pedagogy has long emphasized the teaching of grammar through form-focused instruction. From the perspective of a former student at PUCV and a professional translator, grammar seems to still be the focus for language teaching, though theory offers evidence supporting vocabulary as the aspect teachers should dive into (Thornbury, 2002; Lewis M., 1993, 1997a, 1997b; Nation, 2001).

In a similar view, from the perspective of a professional interpreter who also teaches at PUCV, collocations and the development of collocational competence are essential for future translators and interpreters, as collocational competence is a requirement for communication in the mastery of the language in that one of the objectives of language learners is to achieve native-like English. This idea aligns with Fan's (2009) view that collocational competence is key for producing native-like language. As Hill (2000) points out, learners and language can convey complex ideas efficiently by using collocational expressions, which allow for effective communication.

With regards to the second aspect found within the first category—the importance of collocational competence for translators and interpreters in general and domain-specific translation/interpretation—both teachers and professionals agreed on this. There is in fact evidence in this respect since the appropriate use of collocations in the translation of oral and written discourse is necessary for specialized translation (Baker, 1992). Baker (1992) further explains that speakers who specialize in translation need a working knowledge of genre/register specific collocations to achieve accuracy in translation. It is then the teacher's belief that awareness of word combinations in different languages is of paramount importance in translation since interpreting lexical items accurately implies considering the range of words an item collocates with. Hence, not only is the knowledge of collocations crucial for general translation, but it is also essential for discipline-specific translations which are often required from these professionals (Baker, 1992).

It is very common for translators and interpreters to translate a wide range of genres which belong to specialized areas such as science, technology, history, etc. Translators and interpreters are frequently requested to have experience in translating specialized discourse which exhibits different kinds of collocational features, making these texts more challenging to translate if the professional is not familiar with the range of collocations which are undoubtedly the essence of that piece of discourse). The teachers interviewed expressed that it is of great importance for a translator or interpreter to be competent in this area when translating documents or discourse which is discipline-specific since lexical and grammatical collocations occur in different registers with varied frequency, and this depends on the genre we're looking at, for instance, one possible combination of words might be more frequent in a scientific journal for medicine or science, rather than in a newspaper column, or in the yellow press. Collocations are important in terms of their frequency, and when translating into Spanish or any other language collocations are vitally important.

Further evidence to support this is given by the professional translator interviewed, who expressed that clients look for location services, which means that the product offered needs to sound completely natural within their specialized field of knowledge. The translator explains that the only way to achieve this is through the proper use of collocations. As it has been evidenced, it is almost absurd to envision a discipline-specific translation, such as a medical journal, without paying close attention to the kind of collocations one can meet in the text. Indeed, a translator who fails at appropriately translating collocations might risk losing their own job or the company's client lists.

In the same line, the interpreter interviewed has made emphasis on the importance of collocations for discipline-specific translation and interpretation, by expressing that equivalence in translation and interpretation is attained at pragmatic level, thus, knowledge of the main features and conventions of different genres and professional, academic, language is fundamental. The interpreter adds that each discipline has its own conventions agreed by a discourse community. If the translation ignores such

conventions, such as certain terminology and phraseology, these translations are rejected or not accepted by the community. The functional or communicative purpose of the translation is not attained and the translator-mediated communication fails.

Translations and interpretations can only be accomplished if the terminology is accurate and accepted by a specific discourse community. Translations neglecting discourse-specific lexical features, like collocations, may result in inaccurate lexical renderings, as highlighted by Baker (1992). Considering a reliable translation in specialized discourse, it is imperative then to add that collocational behavior in different genres contributes to the quality of a translation in that it ensures accuracy and conciseness to the translated discourse (Baker, 1992). In this regard, one of the translators interviewed stated that domain-specific languages have their collocational behavior. Thus, the consistent use of vocabulary is more important in specialized languages than in general languages, as it reveals precision and concision–key key features of specialized communication.

The knowledge of collocations is essential for translators and interpreters since those translations which make appropriate use of them sound more natural and professional, and the meaning of a message is conveyed efficiently and successfully. Indeed, collocations are of paramount importance in domain-specific translation in that accuracy is achieved at a semantic and pragmatic level once all features of the language, grammatical and lexical, have been targeted (Baker, 1992). It is clear then that this feature of the language is essential for translation.

The use of corpora has been proven to be ground-breaking in the field of foreign language learning (Bernardini, 2004). This is due to the impact which bringing corpora into the classroom has had in terms of changing from deductive or rule-governed learning to inductive learning. In the light of this, the two teachers, the two translators and the interpreter interviewed in this study suggested that the benefits of using corpora for language learning go beyond the mere use of this tool as a dictionary. The three benefits of using COCA relate to (a) the quality of the language to which students are exposed, (b) the use of corpora to teach domain-specific terminology and (c) the use of corpora for general and domain-specific translation and interpretation.

Regarding the first benefit—the quality of the language to which students are exposed —, the professionals interviewed agreed that one of the most significant advantages of using corpora for language learning has to do with the quality of the stretches of language to which students are exposed. These samples are described as authentic language data which reflect the real use of language in spoken and written discourse. In this light, Granath suggests a series of teaching exercises which aim at exposing students to authentic language patterns through a corpus (Granath, 2009). According to professionals, using a corpus helps learners by providing firsthand experience with authentic language, a key aspect of successful language learning. Moreover, the professionals explain that the issue of frequency is relevant for language learning since this first-hand experience corresponds with the frequency of certain grammatical and lexical patterns of language which learners discover in the use of the corpora. Frequency, then, is of great importance for language learning to take place (Nation, 2001).

The second benefit of using COCA relates to the use of the corpus in the classroom for learning and using terminology, an aspect which was also recognized by the students of this study. As the second translator interviewed for this study is also a teacher in the Translation and Interpretation Program at PUCV, she recognized the importance of using a corpus for teaching terminology, especially since the courses she teaches are part of the

translation/interpretation program. Her stance on using corpora for learning/using terminology corresponds with what research suggests, vocabulary/collocation learning/using is enhanced using different types of corpora (Bernardini, 2004; Romer, 2009).

A third benefit which the professionals recognized in terms of a corpus-based instruction was the use of the corpus for general and domain-specific translation. This was also acknowledged by the students. It seems, therefore, that the use of the COCA is among the corpus software which most translators rely on. The same is also true for domain-specific translations, as one of the translators explains that this is strictly related to using word combinations since the translation memory stores word patterns which are then recovered depending on the type of document to be translated, especially if it is from one specific field, such as medicine.

Along with advantages, the professionals also elicited some negative perceptions about the use of COCA in the language classroom. The first disadvantage, that of time demands, was recognized as the most common problem teachers might face when using COCA in the classroom. This constraint was also identified by the learners who acknowledged that using COCA in class and outside the classroom can be time-consuming. This is perhaps the main reason why teachers avoid using this kind of learning tool in class. Time is very much appreciated when there is a limited number of hours a week students have English lessons (Granath, 2009). Hence, implementing other learning tools which are not mainstream seems to be difficult for some teachers.

The second drawback of using COCA was recognized by one teacher; this has to do with the type of training required for learners to effectively use and take advantage of the learning tool. The teacher, an expert in writing, explains that learners need to know how to interpret the data they find in the interface. Although learners need not be experts in using the software, it is of course necessary for them to have at least some instruction on the different features of the online corpus. In this light, Sinclair (2004) claims that for students to be successful in the use of corpora there is a need to consider their proficiency and experience in using the corpus; this is because it is easy for non-experts to derive wrong conclusions on the evidence provided by the corpus.

It is worth noting that the two drawbacks regarding the use of the COCA which were mentioned in the interviews were identified by teachers of English, and none were stated by the two translators, nor by the interpreter. This is probably due to the need for corpora use in translation and interpretation studies, as opposed to the need to use a wide variety of language strategies and methodologies for general language teaching.

Conclusions

This research was conducted to identify the perception of a group of students from a Translation/Interpretation program regarding their learning and using of collocations through the Corpus of Contemporary American English. It also aimed to identify the perceptions of teachers, translators and interpreters on collocational competence and corpus-based instruction to using and translating collocations. The final aim of this research was to determine to what extent a corpus-based instruction is efficient for the learning and using of collocations as an alternative approach to teaching these lexical items to students from a Translation and Interpretation program.

Results of the study indicate that the students have a positive perception towards learning collocations through a corpus-based instruction, and that they favored using COCA not only to learn, but also to use collocations in different tasks including translation. The data analyzed in this study also revealed that the professionals interviewed believe

collocational competence is necessary for learners to achieve mastery of the target language, and that to accomplish this, a corpus-based instruction would be highly beneficial for them.

Through the analysis of the students' journals and their opinions in the focus group, it was possible to fulfill the first and second objectives of this investigation. On the one hand, it was possible to identify the students' perceptions regarding the learning and using of collocations and, on the other, the extent to which the students perceive corpus-based instruction as appropriate. The analysis of the data revealed 4 different categories which reflect students' awareness of collocations and collocational competence, the reasons why the learning and using of collocations is relevant for them, the benefits of using the Corpus of Contemporary American English for learning and using collocations and the disadvantages of a corpus-based instruction.

Students conveyed that collocational awareness is primarily a university-level skill, not a school-level one, and that learning collocations is highly relevant for language learning, especially translation and interpretation. Finally, students recognized 5 benefits of using COCA: (a) using the corpus for learning collocations, (b) using the corpus to achieve an active role in their learning process, (c) using the corpus to learn specialized terminology, (d) using the corpus to improve the written production of English and (e) using the corpus for translation/interpretation purposes. The only disadvantage to using the corpus which was identified was related to time-constraints, though students believe this can be overcome once they learn how to use all the features of COCA. These findings correspond with the literature in that, on one hand, the Lexical Approach highlights the importance of raising students' awareness of collocational competence and of developing the ability to chunk language successfully (Lewis, 1993). On the other hand, DDL tasks, where learners use corpora by themselves, have proven to be effective to improve vocabulary learning (Cobb, 1999, 2003).

Through the analysis of the teachers, translators and interpreters' interviews it was possible to fulfill the third specific objective of this study: to identify their perceptions on collocational competence and a corpus-based instruction to using and translating collocations. In this respect, 3 categories were found: (a) the importance of teaching collocations and developing collocational competence, (b) the advantages of using COCA, (c) the disadvantages of using COCA. To this regard, it was possible to see that the professionals in the education and translation/interpretation field believe teaching collocations is essential to help students achieve effective communication in the target language, produce native-like English, and learn/use specialized terminology/collocations. Also, the professionals acknowledged several advantages of using COCA, the most mentioned ones being (a) the quality of the language to which students are exposed, (b) the use of this corpus for translation and interpretation of general and domain-specific discourse. Finally, the professionals also identified two possible disadvantages of corpus-based instruction: (a) time constraints and (b) the kind of training required to effectively use COCA. Literature supports the perceptions of these professionals since research has shown that collocational competence is needed in the field of translation/interpretation (Manafi & Ghaffarof, 2013) and that a corpus-based approach to teaching collocations has proven to be effective (Cob, 1999, 2003; Johns, 1991; Bernardini, 2004; Romer, 2009; Granath, 2009).

From what was observed in this study, it is of great importance to consider the teaching of collocations and the development of collocational competence as part of the different English courses in the program. This is due not only to students and professionals' perceptions of collocations, but also to the literature which supports this

approach to language teaching. Based on the results in this study, it is essential to change the approach to teaching these lexical items in terms of the methodology which is currently being used. In this sense, a corpus-based approach to teaching collocations should aim at improving students' lexical proficiency. Moreover, this type of instruction has proven to be appropriate for raising students' awareness of collocational competence; hence, there is a clear need for learners to be instructed in the use of corpora in the language courses as translation and interpretation trainees.

Although the results of this study do point to the need to change the approach towards teaching collocations through corpus-based instruction, there are some limitations which relate to the number of participants. Due to time constraints, only two teachers, two translators, and one interpreter were interviewed to achieve the third specific objective of the research, and this might as well be considered a limitation since a larger sample is needed to validate the findings of this research. However, due to the positive attitudes towards this investigation, we might feel positive in finding similar results if this research were to be extended in terms of the participants.

From this, more inquiries arise for further research: (a) is there a way to motivate learners to use COCA for vocabulary/collocation learning outside the classroom? (b) what is the impact of collocational awareness in oral and written production? (c) what is the impact of collocation knowledge for reading and listening comprehension? and (d) what is the impact of the use of parallel corpora to teach vocabulary to students from a Translation/Interpretation program? This study revealed the perceptions of different actors involved in the Program of translation/interpretation on a corpus-based approach to learning and using collocations. Thus, a more profound study could be done to analyze the impact of collocations on productive and receptive skills and to analyze the use of other corpora for language learning and for translation and interpretation. This would allow a deeper understanding of the importance of collocational competence and the use of corpora for students in Translation/Interpretation studies.

References

Baker, M. (1992). *In other words: A coursebook on translation*. Routledge.

Bernardini, S. (2004). Corpora in the classroom: An overview and some reflections on future developments. En J. Sinclair (Ed.), *How to use corpora in language teaching* (pp. 15-36). John Benjamins.

Cobb, T. (1997). Is there any measurable learning from hands-on concordancing? *System, 25*(3), 301-315. https://doi.org/10.1016/s0346-251x(97)00024-9

Cobb, T. (1999). Breadth and depth of lexical acquisition with hands-on concordancing. *CALL Journal, 12*(4), 345-360. http://dx.doi.org/10.1076/call.12.4.345.5699

Cobb, T. (2003). Do corpus-based electronic dictionaries replace concordancers? En B. Morrison, G. Green, & G. Motteram (Eds.), *Directions in CALL: Experience, experiments, evaluation* (pp. 179-206). Hong Kong Polytechnic University.

Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than non formulaic language by native and non-native speakers? *Applied Linguistics, 29*(1), 72-89.

Davies, M. (2010). The Corpus of Contemporary American English as the first reliable monitor corpus of English. *Literary and Linguistic Computing, 25*(4), 447-464.

Ellis, N. C. (1996). Sequencing in SLA. *Studies in Second Language Acquisition, 18*(1), 91-126.

Ellis, R. (1997). *Second language acquisition*. Oxford University Press.

Ellis, R. (2001). Focusing on form: Towards a research agenda. En W. Renandya & N. Sunga (Eds.), *Language curriculum and instruction in multicultural societies* (pp. 123-144). SEAMEO Regional Language Centre.

Fan, M. (2009). An exploratory study of collocational use by ESL students: A task-based approach. *System, 37*(1), 110-123. https://doi.org/10.1016/j.system.2008.06.004

Firth, J. R. (1957). *Papers in linguistics 1934-1951*. Oxford University Press.

Gass, S., & Selinker, L. (2001). *Second language acquisition: An introductory course*. Lawrence Erlbaum Associates.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Aldine De Gruyter.

Granath, S. (2009). Who benefits from learning how to use corpora? En K. Aijmer (Ed.), *Corpora and language teaching* (pp. 47-65). John Benjamins.

He, M., Xie, Q. Empowering autonomy in language learning: the sustainable impact of data-driven learning on noun collocation acquisition. *Humanit Soc Sci Commun* 11, 1516 (2024). https://doi.org/10.1057/s41599-024-04038-6

Hill, J. (2000). Revising priorities: From grammatical failure to collocational success. En M. Lewis (Ed.), *Teaching collocation: Further developments in the Lexical Approach* (pp. 47-69). Language Teaching Publications.

Horst, M., Cobb, T., & Nicolae, I. (2005). Expanding academic vocabulary with an interactive online database. *Language Learning & Technology, 9*(2), 90-110. http://dx.doi.org/10125/44021

Howarth, P. (1998a). Phraseology and second language proficiency. *Applied Linguistics, 19*(1), 24-44. http://dx.doi.org/10.12691

Howarth, P. (1998b). The phraseology of learners' academic writing. En A. Cowie (Ed.), *Phraseology: Theory, analysis and applications* (pp. 161-186). Clarendon Press.

Johns, T. (1991). Should you be persuaded: Two samples of data-driven learning materials. En T. Johns & P. King (Eds.), *Classroom concordancing* (pp. 1-16). *ELR Journal*, 4.

Johns, T., & King, P. (Eds.). (1991). *Classroom concordancing*. Special issue of *ELR Journal* 4. Centre for English Language Studies, University of Birmingham.

Lewis, M. (1993). *The Lexical Approach: The state of ELT and the way forward*. Language Teaching Publications.

Lewis, M. (1997a). *Implementing the Lexical Approach: Putting theory into practice*. Language Teaching Publications.

Lewis, M. (1997b). Pedagogical implications of the lexical approach. En J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 255-270). Cambridge University Press.

Mackey, A., & Gass, S. M. (2005). *Second language research: Methodology and design*. Lawrence Erlbaum Associates.

Manafi, A. S., & Gh, S. (2013). Effect of collocational competence on translation accuracy of translation trainees. *Journal of Advances in English Language Teaching, 1*(3), 76-84. http://dx.doi.org/10.2139/ssrn.4386585

Mohammed, T. A. S. (2022). *The use of corpora in translation into the second language: A project-based approach*. *Frontiers in Education, 7*, 849056. https://doi.org/10.3389/feduc.2022.849056

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.

Nattinger, J. R., & DeCarrico, J. S. (1992). *Lexical phrases and language teaching.* Oxford University Press.

Richards, J. C., & Rodgers, T. (2001). *Approaches and methods in language teaching.* Cambridge University Press.

Römer, U. (2009). Corpus research and practice: What help do teachers need and what can we offer? En K. Aijmer (Ed.), *Corpora and language teaching* (pp. 83-98). John Benjamins.

Satchayad, P., & Charubusp, S. (2022). The impacts of blended corpus-based instruction on enhancing writing proficiency of Thai university students. *The Asian EFL Journal, 24*(6), 132-156. https://www.asian-efl-journal.com

Shaw, E. M. (2011). *Teaching vocabulary through data-driven learning* (Unpublished master's thesis). University of Texas, US.

Sinclair, J. (Ed.). (2004). *How to use corpora in language teaching.* John Benjamins.

Thornbury, S. (2002). *How to teach vocabulary.* Longman.