

## **Boosting English Vocabulary Knowledge through Corpus-Aided Word Formation Practice**

### **Fomento del conocimiento del vocabulario en inglés a través de una práctica basada en corpus sobre los procesos de formación de palabras**

ANA GONZÁLEZ-MARTÍNEZ

EVELYN GANDÓN-CHAPELA

UNIVERSIDAD DE CANTABRIA

Using a language fluently involves knowing plenty of words and much information about them (Willis, 2003). Native corpora provide an opportunity to access millions of words and their characteristics in a variety of formats through real patterns of vocabulary use (Elgort, 2018). However, there is still a gap between theory and the actual implementation of corpora in the classroom (Römer, 2006). This paper extends previous works focused on learning through corpora in different educational levels, such as the activities suggested by Roca Varela (2012), since other examples of coursework including direct native corpora use in an English learning context are scarce outside the university level (for example, see Matos, 2013). In this paper we propose a sequence of activities to promote morphological awareness by taking a closer look at the diverse processes of word formation of the English language through the COCA (Davies, 2008-) and BNC (Davies, 2004) corpora within the Spanish Secondary Education context.

**Keywords:** *corpora; EFL teaching; corpus work; word formation; vocabulary teaching and learning*

El uso fluido de una lengua supone conocer un gran número de palabras y una amplia información sobre ellas (Willis, 2003). Los corpus nativos brindan la oportunidad de acceder a millones de palabras y sus características en diversos formatos con ejemplos reales de uso del vocabulario (Elgort, 2018). Sin embargo, aún existe un largo camino entre la teoría y el verdadero uso directo de los corpus en el aula (Römer, 2006). En este artículo se realiza una propuesta de implementación práctica en Educación Secundaria como la presentada por Roca Varela (2012), ya que otros ejemplos en niveles educativos fuera del contexto de la educación universitaria son escasos (véase Matos, 2013). Además, se proponen algunas actividades que promueven la conciencia morfológica, analizando diversos procesos de formación de palabras en inglés a través de los corpus COCA (Davies, 2008-) y BNC (Davies, 2004) dentro del marco de la Educación Secundaria española.

**Palabras clave:** *corpus; enseñanza del inglés como lengua extranjera; trabajo de corpus; formación de palabras; enseñanza y aprendizaje de vocabulario*

## **1. INTRODUCTION**

As is well-known, the concept of *corpus* refers to an electronic collection of texts. Thanks to these computer-based corpora, researchers can access data as never before in terms of quantity

and quality (Sinclair, 1999: 1). It could be considered that corpus linguistics has completely changed the landscape of language study, as their typology and applications have extended to language instruction as well. In particular, native electronic corpora are potentially a great asset for vocabulary learning. Generally, it has been accepted that acquiring a solid vocabulary is essential for every step in the language learning process that aims for communication (Canale & Swain, 1980), but it is now known that word learning goes beyond amount (Nation, 2000: 26). Caro and Mendinueta (2017: 208) establish that there is a contrast between the concepts of *vocabulary breadth* and *vocabulary depth*, where the former refers to the fact of knowing many words and the latter to that of knowing diverse aspects of them. Although it is accepted that paying attention to word formation is an important vocabulary learning strategy that helps learners with meaning retention (Nation, 2000: 264), numerous teachers assume that these processes do not require explicit teaching, because they are assumed to be inferred mechanically as the learner progresses (Tahaine, 2012: 1106). Nevertheless, this is not the case for all learners, and many of them will acquire incomplete vocabulary depth, which may hinder their competence in the target language. Through corpora, learners have the opportunity of approaching word learning in context (Ma & Kelly, 2006: 16), which could help with the familiarisation of the multiple dimensions of a particular word, providing information such as its meaning or use. Yet, due to factors such as teacher unawareness or lack of training, corpora are still a long way from finding their place inside the language classroom. Authors like Roca Varela (2012) have suggested how to exploit corpora to practise English vocabulary at different academic levels. However, other examples of coursework including direct native corpora use in an English learning context are scarce outside the context of Tertiary Education and English for Specific Purposes (Gabrielatos, 2005: 5) (see, for example, Matos, 2013). For this reason, the aim of this paper is to illustrate how native electronic corpora could be applied directly in an English as a Foreign Language (EFL henceforth) classroom, based on the context of Secondary Education within the Spanish curricular framework. The activities will focus on the study of morphology through two native corpora: the Corpus of Contemporary American English (COCA, 2008-) by Mark Davies and the British National Corpus (BNC, 2004) by Oxford University Press<sup>1</sup>. In these activities, corpora are the foundation for the study of words and their formation processes, vocabulary, and language learning. Materials, timing, and assessment guidance are also suggested in order to overcome the drawbacks that direct corpora application in the classroom may cause. Five sample sessions are provided to get the students acquainted with the concept and functions of the BNC (Davies, 2004) and COCA (Davies, 2008-) corpora so that they can experience a hands-on approach to different word formation processes.

The paper is structured as follows: Section 2 offers a theoretical review of the influence of corpus linguistics on the realm of language teaching and learning. This part analyses the beneficial aspects and challenges that corpora bring into the classroom, and pinpoints which matters need to be addressed for a more effective teaching practice. Additionally, this section emphasises the importance of word formation processes and their impact on word knowledge. Section 3 targets the aforementioned challenges, the methodology, and other practical considerations that need to be accounted for in order to move from theory to the actual implementation of direct corpora use in the classroom. Finally, Section 4 presents a practical proposal that illustrates how corpora can be implemented in a Secondary Education EFL class to promote vocabulary learning and morphological awareness through word formation practice.

---

<sup>1</sup> These corpora can be accessed at <https://www.english-corpora.org/coca/> and <https://www.english-corpora.org/bnc/>

## 2. NATIVE ELECTRONIC CORPORA IN THE LANGUAGE CLASSROOM

In this section, we will offer an overview of the benefits and challenges brought by the direct application of native electronic corpora in the second language classroom. We will also tackle the issue of what knowing a word implies, and the role of word formation for this purpose. Finally, we will analyse the importance of a thorough acquisition of vocabulary for the correct development of a second language, with special emphasis on its vital role in becoming a proficient user of such a language.

Due to the diffusion of corpus linguistics, corpora have also reached the realm of language teaching and learning. In the language classroom, native corpora have been applied indirectly and directly (Römer, 2011: 207). The indirect approach places the focus on researchers, who use corpus evidence to examine language in use and to study how corpora may contribute in a learning environment (Römer, 2011: 206). Römer (2006) and Conrad and Levelle (2008) distinguish different types of indirect pedagogical corpus applications, which include using corpora to improve course designs and prepare class syllabi and materials, like dictionaries. In contrast, following a direct approach implies that “teachers and learners get their hands on corpus data themselves, instead of having to rely on the researcher as mediator or provider of corpus-based materials” (Römer, 2006: 124). This implies that learners themselves perform corpus searches to acquire linguistic knowledge about a particular language, thus opening up new possibilities for teachers and students. These electronic collections of words may contribute to the development of the L2 (Gabrielatos, 2005: 20), but they can also bring new challenges. Such matters will be examined in further detail in the following subsections.

### 2.1 *The benefits of learning through corpora*

Above all, electronic corpora are considered to be authentic, since they allow examining naturally-occurring language data that is produced in real communication situations (Gilquin & Granger, 2010: 2). Further, they provide variety with a large number of samples of a particular item that can be studied in different contexts and frequencies (Gilquin & Granger, 2010: 2; Gabrielatos, 2005: 14). According to Asención-Delaney, Joseph, Collentine, Colmenares and Plonsky (2015: 143), learning through corpora provides learners with multiple vocabulary use samples through a wide variety of concordance lines.

Another particular benefit for language learners is that corpora may be autonomy-promoting and particularly adequate for learning lexis, as claimed by Poole (2018). As a result of using corpora, students have more freedom and become more responsible for their own instruction (Gilquin & Granger, 2010: 5). In fact, Conrad and Levelle (2008: 548) observed that learner autonomy increases as students learn how to make generalisations based on observable data, instead of relying completely on the knowledge presented by their teachers.

Moreover, corpora may also be an important motivational element in the acquisition of a particular language. This is due to the fact that following an inductive approach may be appealing for those students with different learning styles or needs, instead of the traditional approach based on language rules (Conrad & Levelle, 2008: 548). Corpora may enhance the discovery factor of learning, with students taking the role of language researchers (Gabrielatos, 2005: 20). Finally, corpora are considered innovative, as learners explore language while incorporating the use of new technologies (Gabrielatos, 2005: 20).

## 2.2 *The drawbacks of learning through corpora*

Learning through corpora also imposes some limitations concerning its direct application in the classroom. As Asención-Delaney et al. (2015: 141) claim, a limited number of studies have measured lexical development through corpora, and most research is focused on student perceptions about the use of corpora as a method for language learning.

One of the main obstacles a language teacher may find is that creating corpus-based lessons may be costly in terms of resources. From a material point of view, at least one computer will be needed for every pair of students, together with access to corpora and other software. All this costs money, and some schools are not always able to afford them (Gilquin & Granger, 2010: 7). Furthermore, Gilquin and Granger (2010: 7) point out that even though some corpora are free, they may have more limited features than those bought. Additionally, it is time-consuming to prepare teaching materials, train students in the use of corpora, and complete a search task (Gilquin & Granger, 2010: 7). Lee and Lin (2019: 15) add that students who are less accustomed to inductive learning methods may require more time to make inferences by themselves.

Teacher reticence is another major impediment to the use of corpora. Meunier (2011: 461) blames this on their lack of awareness of the benefits that corpora may provide, while Conrad and Levelle (2008: 548) suggest that there are few empirical studies that shed light on what activities or skills improve the most under this approach, such as Gaskell and Cobb's (2004) study on correcting vocabulary errors in writing tasks. Gilquin and Granger (2010: 2) consider that teachers are not well trained in this field and do not know enough about corpora to aid their students. A knowledge foundation, time, and basic training are therefore essential in order to work with corpora (Römer, 2006). In addition, teachers would need to face and overcome some challenges for the method to succeed, like considering how these materials can be integrated into the curriculum (Breyer, 2009: 156).

Concerning perceived difficulties, Asención-Delaney et al. (2015: 142) state that concordance lines could be difficult to interpret, as the context that is usually provided with them tends to be shortened. Further, Gilquin and Granger (2010: 4) highlight that learners may struggle with some corpus functions, like annotations, the Keyword in Context (KWIC hereinafter) view<sup>2</sup>, or discerning the irrelevant hits. In addition, corpora interfaces may be too sophisticated for novice users (Asención-Delaney et al., 2015: 148).

It has also been considered that learning through corpora may not be suitable for the whole class and not appealing for all students, since some of them may not feel comfortable working with technologies for language learning (Ma & Kelly, 2006: 16). Moreover, some exercises may exhaust the cognitive resources of learners, like their attention, if they find no connection between what they are doing and a context (Asención-Delaney et al., 2015: 142).

Additionally, the amount of autonomy assigned to learners might be unfavourable: too much freedom may affect the learning outcomes (Ma & Kelly, 2006: 16). As Römer (2011: 215) argues, even the complexity of the data may intimidate learners, especially those who still have a limited vocabulary. Further, there may be some students who prefer a more explicit approach to learning (Asención-Delaney et al., 2015: 148)

Lastly, one of the arguments encountered against the use of corpora for language learning is that they oppose a communicative language approach. Leńko-Szymańska and Boulton (2015: 4) believe that a corpus analysis of language is incompatible with a communicative

---

<sup>2</sup>The Keyword in Context (KWIC) view in a corpus shows “the patterns in which a word occurs, by sorting the words to the left and/or right” (Davies, 2008-). Each word in the text is labelled with a colour code (e.g. nouns in blue or verbs in pink).

language teaching methodology because it is an approach that aims for accuracy rather than fluency.

Teachers are beginning to become more aware of the possibilities of using corpora, but there is still a gap between theory and actual pedagogical implementation and a long way to go to shorten that distance (Römer, 2006: 129; McCarthy, 2008: 572). Corpus-based instruction seems to provide multiple benefits for language learners who are still developing the interlanguage (Selinker, 1972). Still, it seems to be necessary to justify this by looking at the evidence provided by research to date. What may be concluded so far is that native corpora have exerted an influence on language education that cannot be ignored.

### 2.3 Word formation and word knowledge

Tahaineh (2012: 1105) defines word formation as the processes involved in “the creation of new words on the basis of existing ones”. The study of the nature of words involves, among other features, the different processes by which terms are formed. Tahaineh (2012: 1108) has established a classification of various word formation processes in English, which suggests that there are recognisable and predictable patterns involved in word building. Some of the most common processes described by this author are the following:

- 1) Compounding: Two or more roots and bases that are joined to produce a new single one, e.g. *handbag* (*hand* + *bag*).
- 2) Borrowing: Loanwords that are borrowed from other languages, e.g. *bazaar* from Persian, meaning *market*.
- 3) Conversion or zero derivation: A lexical item is changed from one grammatical class to another without affixation, e.g. the noun *bottle* (i.e. I bought a *bottle* of soda) to the verb *to bottle* (i.e. Water is *bottled* in the factory).
- 4) Stress shift: When pronounced, the word stress is moved from one syllable to another, e.g. *transport* (/ˈtrænsˌpɔːrt/) to *transport* (/trænsˈpɔːrt/), changing the grammatical class of the word (noun and verb, respectively).
- 5) Clipping: Words of more than one syllable are reduced in casual speech, e.g. *flu* from *influenza*.
- 6) Acronym formation: Terms are formed from the initials of a group of words, e.g. *NASA* (*National Aeronautics and Space Administration*).
- 7) Blending: Two parts of already-formed words are joined to create a new one, e.g. *brunch* (*breakfast* and *lunch*).
- 8) Backformation: A (pseudo-) suffix is removed from the base, and this base is used as a word (e.g. *babysit* from *babysitter* or *burger* from *hamburger*).
- 9) Coinage: Invention of brand-new terms, most of them from a company’s product that becomes the generalisation, e.g. *Kleenex* for *tissue*.
- 10) Onomatopoeia: Words that sound like the sound they name, e.g. *buzz* or *crack*.

- 11) Derivation: It consists in joining affixes and already existing words together to create new terms that belong to a different grammatical category. Some examples of these processes could be forming the noun *direction* from the verb *direct*; forming the verb *shorten* from the adjective *short*; forming the adjective *beautiful* from the noun *beauty*, or forming the adverb *completely* from the adjective *complete*.
- 12) Affixation: Affixation consists in combining affixes with roots, changing the meaning. Some examples of this process are *co-*, as in *co-owner*; *un-* as in *undo*, or *dis-* as in *dishonest*.

Although it is important to understand these processes in order to familiarise oneself with English vocabulary, many language teachers assume that these are not in need of explicit learning, because students will end up, at some point, inferring them (Tahaine, 2012: 1106). However, teaching these mechanisms is an area worthy of attention. Nation (2000: 264) claims that focusing on word parts and word formation processes is a useful strategy for learning new terms because students are more likely to identify affixes and interpret the meaning of the whole word. In other words, this ability contributes to the promotion of word knowledge (Nation, 2000: 270).

The notion of word knowledge leads us back to the matter that words are not independent units with a single dimension. There are many aspects to know and many degrees of knowing any given word (Nation, 2000: 23), and using the language fluently depends on both knowing plenty of words and much information about them (Willis, 2003: 13). As mentioned in Section 1, a distinction is usually made between two dimensions of word knowledge: *breadth* and *depth*.

Nonetheless, knowing a word involves much more than knowing a lot or knowing how it is spelt or pronounced; there are multiple dimensions to recognise, referred to as *vocabulary depth* (Caro & Mendinueta, 2017: 209). For clarification, let us take the word *bubbly* and examine it in a similar fashion to Nation's (2000: 41) analysis of the word *underdeveloped*. According to this classification, knowing *bubbly* implies:

- a) Recognising it when it is heard and producing it with correct pronunciation, including its stress /'bʌbli/.
- b) Familiarising with the written form. This involves recognising it when reading and spelling it correctly when writing.
- c) Accepting that it is built by the parts *bubble* and *-y*, adding them, and being able to relate these parts to their meaning.
- d) Knowing that *bubbly* signals a particular meaning and being able to express it. It can take the form of an adjective, referring to a drink that is full of or produces bubbles, or describe a lively and cheerful person. On the other hand, it can take the form of a noun to refer to *champagne*.
- e) Knowing what the term means in the particular context in which it occurs and producing it with the intended meaning (e.g. as an adjective, referring to an object or a person, or as a noun).
- f) Knowing that there are related words like *fizzy*, *effervescent* or *energetic*, and being able to produce synonyms and opposites such as *still* or *apathetic*.

- g) Recognising the correct use of the word in a sentence and using it appropriately when producing an original one.
- h) Identifying terms such as *personality*, *water* and *bottle* as typical collocations, and producing words that commonly occur near them.
- i) Knowing that bubbly is not an uncommon or pejorative word, and adapting the term to the degree of formality of the situation, knowing that the noun bubbly referring to champagne is an informal use.

Knowing a word is the result of a process that learners have to undergo (Bogaards, 2001: 325). This process implies that before knowing a particular word, learners have to become familiar with it in different contexts (Bogaards, 2001: 327). Therefore, teachers must ensure that learners are presented with vocabulary in a variety of situations and forms. In addition, educators need to become aware of their students' current lexical knowledge to provide the best instructional decisions (Caro & Mendinueta, 2017: 207).

#### 2.4 Vocabulary learning through corpora

Over the past few years, there has been a shift towards more elaborate new vocabulary learning proposals focused on words that leave behind the notion that vocabulary is learnt automatically and unconsciously (San Mateo-Valdehita, 2013: 17). Although the efficiency of corpora or concordances for vocabulary learning has been a disputed issue, they offer a wide spectrum of possibilities for the study of vocabulary. As psycholinguistic research has proved, language processing is sensitive to the frequency of usage and statistical knowledge (Ellis, 2015: 5), and corpora may be helpful in indicating which forms occur more frequently in a variety of contexts. Ma et al. (2006: 24) state that vocabulary is accessed in context instead of being presented in isolation when it is studied through corpora. Thanks to options like KWICs, learners might be able to examine facts about words that are not usually accessible, such as semantic relations, conceptual fields, or collocations. Nevertheless, learning through corpora has brought in new troubles for teachers to be aware of (Pérez-Paredes, Sánchez-Tornel, Alcaraz Calero & Jiménez, 2011: 1), which is why the challenges described in Subsection 2.1.2 need to be addressed.

Sections 3 and 4 showcase how these challenges may be overcome through a learning proposal with English native corpora. The training sessions illustrate how this software may contribute to the promotion of word knowledge by working with different word formation processes.

### 3. PRE-INTERVENTION CONSIDERATIONS

Before translating a learning proposal from theory into practice, four aspects need to be addressed. In this section, we examine the methodological principles underlying the learning process (Subsection 3.1); the challenges posed by the chosen method (Subsection 3.2); how the learning proposal aligns with the curricular framework (Subsection 3.3), and how students' performance will be assessed (Subsection 3.4).

### 3.1 *Methodological principles*

The potential of the direct use of native corpora in the second language classroom has drawn increasing attention in the past few years towards Data-Driven Learning (DDL henceforth). DDL is defined by its coiner, Johns (1991: 1), as a computer-based approach to language learning in which students “discover the foreign language”. In this approach the language learner is the protagonist and turns into a researcher, deriving knowledge through access to linguistic data (Johns, 1991: 1). The role of the concordance is not to provide answers about the language per se, but to provide inferable data that learners can interpret.

To address the matter of working with an inductive approach which may discourage beginners and more teacher-centred students, Lee and Lin (2019: 24) suggest combining DDL with existing or more traditional teaching approaches to reduce the cognitive load involved. These authors claim that both inductive and deductive approaches entail different methods of reasoning, equally effective in fostering vocabulary acquisition and retention. A mixed approach implies that teachers may, at times, step back on their role of traditional instructors, and act as guides for students in their use of corpora. Teachers could be in charge of managing timing in the classroom, confirming the rules examined, and directing their group debates. Furthermore, teachers may aid students in their corpora searches. This way, learners would benefit from both methods.

### 3.2 *Practical considerations*

In this section, we address four essential factors for the correct implementation of the use of corpora in the classroom. These practical considerations involve reducing the costs of using corpora (Subsection 3.2.1) and taking into account that students may not have any previous knowledge of corpora (Subsection 3.2.2). Likewise, bearing in mind students’ individual differences is also crucial (Subsection 3.2.3), along with the need of aiming for communication (Subsection 3.2.4).

#### 3.2.1 *Reducing the costs*

One of the main concerns of implementing corpora in a language classroom is that they can be costly in terms of timing and material resources. Consequently, it is necessary to explore how these two concerns can be effectively addressed.

- a) **Timing.** Direct corpus use may be implemented with flexibility. It may take the form of an exercise with concordance lines (Conrad & Levelle, 2008: 547), a sequence of them or a learning unit, all of which are compatible with other subject matter contents. For instance, a group of students that takes three EFL sessions of 60 minutes per week uses 10% of the weekly study load. If a learning unit is carried out throughout one scholar term (September through December), and only one weekly hour is devoted to a corpus session, it allows for nearly 15 sessions in total that could be used for implementing this approach.
- b) **Material resources.** The use of technological devices is essential for the correct development of these activities, since consultation through electronic corpora requires available devices and Internet access. Despite this, the choice concerning the devices may depend on the resources available at each particular institution. Corpora can be accessed through computers, tablets, or smartphones, and students may share their



devices in pairs. In case no Internet connection is available, teachers may adapt their activities by providing students with result lists extracted and printed out from the corpora. Further, students will need access to corpora. Consequently, we advocate the use of accessible corpora such as the two native corpora chosen for our proposal, the Corpus of Contemporary American English (COCA, 2008-) by Mark Davies and the British National Corpus (BNC, 2004) by Oxford University Press. These two corpora have a simple interface and contain a large amount of authentic native speaker data. Both corpora represent different varieties of English so students may critically analyse language based on parameters like usage, form or adequacy. In terms of costs, they are free and available online (only previous registration through email is required).

### 3.2.2 *Previous knowledge*

Even though it may be assumed that students in Secondary Education have been in contact with ICT resources, it is possible that some have never approached electronic corpora nor encountered the concept of corpus linguistics. Considering that it could be the first time that learners use electronic corpora, one or two sessions may be devoted to ensuring that students understand how corpus linguistics works. This could be carried out prior to engaging with the rest of the activities, so that learners acquire basic corpus search skills and become familiar with language data analysis.

The fact that students may not have enough experience with a mixed approach must also be considered. Thus, in order to aid them, activities should be designed so that students have enough support to carry them out. For this reason, the activity sequence could be arranged following a step-by-step structure, especially during their first experience with the corpora. In addition, the use of reference tools (e.g., dictionaries) besides the electronic corpora would be highly recommended to have an extra aid for meaning consultations.

### 3.2.3 *Individual differences*

As mentioned earlier, one must consider that individual differences and learning styles may be present within the students' group. Concerning students' individual differences, there are some factors that may influence how they acquire the L2. These include personality traits like extroversion or introversion, the level of anxiety towards the L2, or their attitude (Dewaele, 2009). On the other hand, some authors emphasise that the students' learning styles and strategies may boost or withhold a particular methodology. Oxford (2003) claims that if there is harmony between the preferred learning styles and strategies and the methodology, students are likely to feel more confident and, in consequence, perform better. The opposite, Oxford (2003) states, may lead to poor performance and the discouragement of students.

To address the matter of working with an inductive approach which may discourage beginner and more teacher-centred students, Lee and Lin (2019: 24) suggest combining DDL with traditional teaching approaches to reduce the cognitive load. Through DDL and inductive work, students become protagonists of their own learning process. This implies that they become aware of the language features studied and they gain autonomy while task engagement is promoted. Through traditional work, students that are more accustomed to teacher-oriented methods will feel more comfortable, while reducing the difficulty and less positive aspects involved in inductive learning.

### 3.2.4 *Aiming for communication*

Canale & Swain (1980: 9) defined a theory of basic communication skills as "one that emphasizes the minimum level of (mainly oral) communication skills needed to get along, or cope with, the most common second language situations the learner is likely to face". Past work in second language acquisition research carried out by Canale & Swain (1980) suggested that

communicative approaches to language teaching relied more on being understood, that is, meaning, than accuracy. Nonetheless, these initial views soon encountered difficulties. Long, in 1991, stated that a theory of Focus on Form (FonF) consisted in “drawing students’ attention to linguistic elements as they arise incidentally in lessons whose overriding focus is on meaning or communication” (Long, 1991. 45-46, as cited in Laufer, 2005: 224). This approach argued that focusing exclusively on meaning could not help learners achieve the desired level of grammatical competence in the target language, and thus, it would be necessary to pay attention to form as well (Laufer, 2005: 224). According to Laufer (2005: 224), the ideal situation would be to focus on form in a communicative task environment.

In a corpus-led session, the reading material and corpora hits will be the main vehicle for learning, which will lead students to investigate in the corpora. As these texts are written, it would be necessary to design activities in which more than one language skill is practised. In order to promote a communicative model, it would also be positive to select the reading material in a varied and rich way. The texts may be selected according to their genre, vocabulary variety, features of interest, size, and level of complexity. The purpose of this is to ensure not only authentic input, but also that the texts are rich in terms of language content, adequate for the language level of the students and suitable in terms of size.

### *3.3 Establishing a link with the curricular guidelines*

First, it is necessary to establish the framework on which the activities are based. Understanding the educational background in which a learning unit is to be carried out is necessary to understand the implications of bringing corpora into the language classroom within this context.

For the purposes of our proposal, the document used for reference is the Royal Decree 1105/2014 of December 26<sup>th</sup>, whereby the basic curricula of Secondary Education and Baccalaureate are established (2014). This decree establishes the curricular guidelines for Secondary Education in Spain. The First Foreign Language subject (usually English) is integrated in the curriculum as a basic subject in learners’ formation, and it is grounded in the Common European Framework of Reference for Languages (Council of Europe, 2001). Learners are accordingly expected to be able to apply the acquired knowledge and skills in real interaction processes, with communication as the final purpose.

The subject is divided into four main blocks according to each communicative skill: oral comprehension (listening), oral production (speaking), written comprehension (reading), and written production (writing). Each of these blocks presents the contents, assessment criteria and learning outcomes necessary for each stage. At every level of Secondary Education, the amount of lexis, the different aspects of lexis that must be known or the specific lexical items that must be taught are not explicitly stated. Guidance about how to proceed with the teaching of lexis is not explicitly stated either. Only the common topics of vocabulary and the fact that it must be recognised and used properly are specified.

Although in this curriculum lexis is embedded within other competences, it still holds great importance in the communicative situation, as all tasks involving meaning, comprehension or inferring place a lot of weight on lexis. Furthermore, those tasks that require using adjectives, nouns, adverbs or verb conjugations are related to morphology. Nevertheless, the fact that these contents are not specified gives teachers the freedom to select those contents that they consider necessary. However, this could also be a problem. In many cases, leaving an open choice may lead to a wide difference in lexis knowledge among groups of students. For instance, one teacher may consider studying affixation necessary while another may not. As a result, the amount, the knowledge of different word aspects and the strategies students know

and use to cope with gaps in their vocabulary may vary greatly at this stage of language learning.

With the study of morphology through native corpora, students will become involved in some of the competences that were established in the curriculum. For instance, apart from the main competences of linguistic communication and digital competence, learners will be targeting the mathematical and scientific competences by working with the language data offered by the corpora and by creating inferences and hypotheses.

### *3.4 Assessing performance*

Once the sessions have been carried out, it is necessary to evaluate the students' learning process. The assessment may be carried out by both students and teachers. Student self-assessment may be helpful for teachers to deal with aspects such as the lack of time to assess every student in large classrooms (Jamrus & Razali, 2019: 71). Further, it is an encouraging technique that enables students to become more autonomous, i.e. active learners (Vasu, Mei Fung, Nimehchisalem & Rashid, 2020).

In particular, portfolio-based assessment plays an indispensable role in language self-assessment. Kohonen (2000: 1) points out that many aspects of learning a language can only be inferred in an indirect way based on the output produced by students. This knowledge may be unconscious, remaining out of reach for teachers and students to address. Nevertheless, portfolios may be a visual representation of the development of the students' learning process (Ma'arif, Abdullah, Fatimah & Hidayati, 2021: 8). They offer teachers the possibility of helping their students to become more aware of the learning goals and outcomes (Kohonen, 2000: 2). As an example, a learner portfolio in a corpus-guided lesson may serve as an assessment tool and as a classroom diary in which students will record the key elements to remember, the language features studied and their reflections, thoughts and attitudes on the lesson. This way, the teacher and stakeholders of the particular institution will be able to observe students' development, as well as their difficulties and needs. It will also help students by promoting critical judgement of their own work and by raising self-consciousness. This will help address any difficulties that may occur during the implementation of the corpus activities.

## **4. A SPECIFIC PROPOSAL OF IMPLEMENTATION IN SECONDARY EDUCATION**

### *4.1 Context*

As this proposal has not been implemented in a classroom thus far, this learning unit has been designed for a sample group of 16 students (aged 17) of the subject of English as First Foreign Language in the educational stage of 2<sup>nd</sup> grade of Baccalaureate in a state-funded high school in Spain. They have been enrolled in a Content and Language Integrated Learning (CLIL) programme from the beginning of Compulsory Secondary Education (age 12), and they have been studying EFL from the beginning of Primary Education (age 6). Throughout these school years, they have all been in contact with conversation assistants from different countries, and most of them have participated in abroad programmes offered by their school or their extracurricular language centres.

Regarding the materials available, the educational centre has an ICT room equipped with 25 computers. Furthermore, tablets and laptops are available for student loaning at the school library, in case students need them for personal study. All sessions would be carried out at the

centre, so no extracurricular time is needed. Nevertheless, students are encouraged to practise on their own and research.

The proposal provided here corresponds to five sample sessions of 60 minutes, arranged as follows: one session devoted to getting acquainted with the notion of electronic native corpora, three sessions for the study of different word formation processes, and one review session.

## 4.2 Aims

The purpose of this paper is to provide an example, from a pedagogical point of view, of how the concept of corpus could be introduced in the EFL classroom, and how different word formation processes may be investigated using native electronic corpora to facilitate vocabulary learning.

## 4.3 The activities

### 4.3.1 Getting the concept of corpus linguistics and basic training (Session 1)

- a) Manual corpora (20'): The main purpose of this exercise is for students to comprehend the notion of corpus linguistics, as it may be their first contact with corpora. To fulfil this aim, two text fragments will be handed out to students, who will work in pairs. Learners will have to read the texts and highlight, using colours, a word that is repeated in both texts (Appendix 1). Students will then have to count the number of hits of this word and analyse it in terms of frequency, the part of speech it belongs to, and suggest some possible collocations and synonyms. After performing this task, the teacher will reveal a faster way to do all this, by introducing the notions of electronic corpora and corpus linguistics.
- b) Guided search in electronic corpora (15'): With the teacher's guidance, and using the worksheet provided in Appendix 2, students will conduct a guided search in the established corpora. In this search, learners will explore the basic features of a corpus search (e.g. list, chart, collocates, or compare<sup>3</sup>) and answer a series of questions.
- c) Autonomous search practice (10'): To apply the knowledge about corpora acquired in the previous activity, students will be invited to perform a search on a term of their preference with regard to the functions examined in the previous exercise.
- d) Portfolio work (5'): This time will be devoted to working on the personal portfolio. Students will record the knowledge they have learnt and reflect on their practice and attitude.

### 4.3.2 Word formation processes (Sessions 2, 3 and 4)

Three sessions are proposed in this section. They all follow the same structure, except for the "practice" activity. It is also important to note that the topic of word formation would have been previously introduced in other lessons. The structure proposed consists of the following sequence:

---

<sup>3</sup> The 'List' function shows the frequency and contexts in which the word/phrase appears. 'Chart' performs a term search comparing its frequency in each genre section. 'Collocates' allows observing which words occur more frequently next to another. 'Compare' allows comparing two terms to identify a pattern of occurrence.

- 1) Introduction and organization (5'-10'): Teachers activate students' prior knowledge about the topic by asking questions (e.g. "What parts can you recognise in the word *cooperation*?"). They also inform about the class' structure and timing.
- 2) Text analysis (10') (Appendix 3): A text is presented to students, who select and classify the target vocabulary. This vocabulary is signalled (underlined) by the teacher beforehand. The text and terms that are analysed vary depending on each featured word formation process that is being studied. In this proposal, we have selected affixation, derivation, and onomatopoeias for illustration purposes.
- 3) Mind map (20') (Appendix 4): Students are asked to complete a mind map based on their predictions and hypotheses and then check them on the corpora. They will analyse a particular root or affix by exploring aspects such as meaning, collocations, variant and register differences, part of speech, pronunciation, topics (words that co-occur on the same page as the target term) or clusters (the most frequent word strings). This includes looking up words with the same affixes (e.g. *anti*\* for the prefix *anti-* as in *antibiotic*) (see Figure 1) in COCA (Davies, 2008-) and BNC (Davies, 2004). Additionally, they will have to examine the occurrences using different functions of the corpora (Figures 2 and 3) and, if necessary, consult a dictionary as an extra aid for issues concerning meaning and pronunciation<sup>4</sup>.

ON CLICK: [CONTEXT](#) [TRANSLATE \(??\)](#) [GOOGLE](#) [IMAGE](#) [PRON/VIDEO](#) [BOOK](#) (HELP)

HELP		ALL FORMS (SAMPLE): 100 200 500	FREQ	TOTAL 211,293   UNIQUE 11,096
1	<input type="checkbox"/>	ANTICIPATED	11877	
2	<input type="checkbox"/>	ANTIQUÉ	7405	
3	<input type="checkbox"/>	ANTIBIOTICS	7320	
4	<input type="checkbox"/>	ANTICIPATION	7010	
5	<input type="checkbox"/>	ANTICIPATE	6900	
6	<input type="checkbox"/>	ANTIQUES	4370	
7	<input type="checkbox"/>	ANTIBODIES	3530	
8	<input type="checkbox"/>	ANTIBIOTIC	3519	
9	<input type="checkbox"/>	ANTICIPATING	3464	
10	<input type="checkbox"/>	ANTITRUST	3447	

Figure 1: Example of a LIST search in COCA concerning the prefix *anti-*

CLICK FOR MORE CONTEXT NEW SAVE

1	2012	BLOG	addinginfo.org	Q	... Accomplishments! With Citations! " was published. # FDA ordered to reconsider <b>antibiotic</b> regulations # A US federa
2	2012	BLOG	...dbelord.blogspot.com	Q	tried drinking soup but could barely take a few spoons. I immediately took my <b>antibiotic</b> and went to sleep. Had to wal
3	2012	BLOG	mezorex.com	Q	3254406 The assisted living arm of the American Health Care Association satisfied cipro <b>antibiotic</b> price. # ALFA is the
4	2012	BLOG	mezorex.com	Q	last year, including the Alzheimer's Association, Center for Medicare Advocacy Inc cipro <b>antibiotic</b> price. National Comr
5	2012	BLOG	mezorex.com	Q	of a genetic test before surgery to determine the risk of patient s. # <b>antibiotic</b> price # E-mail calling about ramped fight
6	2012	BLOG	...hecenter.typepad.com	Q	Every morning I'd get up, take a big old syringe filled with yellow <b>antibiotic</b> out of the fridge and place it on the dining r
7	2012	BLOG	...hecenter.typepad.com	Q	my left bicep starting at the left elbow with a saline syringe, administering the <b>antibiotic</b> syringe, followed up with a an

Figure 2: Example of a LIST search in COCA concerning the occurrences of *antibiotic*

<sup>4</sup> E.g. Cambridge Dictionary, which provides the phonetic transcription of the word <https://dictionary.cambridge.org/>

**squeak** (VERB) #13125

See also: **NOUN**

BLOG WEB TV/M SPOK FIC MAG NEWS ACAD

1. make a high-pitched, screeching noise, as of a door

DMOCGE

YouGlish PlayPhrase Yarn

Translate: choose language

**SYNONYMS** NEW: DEFIN +SPEC +GENL  
squeal, screech, squeak, squeal, whine, yelp

**TOPICS** (more)  
fur, lick, grin, pat, squeak, tighten, kneel, softly, gaze, click, mutter, refrigerator, robe, scream, breeze, cough, poke, straighten, faint, frown

**COLLOCATES** (more)  
NOUN door, chair, shoe, voice, floor, wheel, victory, win  
VERB manage, rattle, squeal, groan, bounce, rock, squeeze, pop  
ADJ open, wet, rubber, front, loud, rusty, wooden, narrow  
ADV by, barely, through, loudly, somehow, twice, forth, open

**RELATED WORDS**  
squeaky, squeaker, squeaking, squeaking, pipsqueak

**CLUSTERS** (more)

squeak •	squeak out • squeaked out • squeaked by • squeak by • squeaked in • squeaked by • squeaking on • squeaked on
• squeak	he squeaked • it squeaked • door squeaked • barely squeaked • voice squeaked • she squeaked • i squeaked • chair squeaked
squeak ••	squeaked by with • squeaked as he • squeaked as she • squeaked when he • squeaked by in • squeaked as i • squeak when they • squeak when you
•• squeak	managed to squeak • screen door squeaked • able to squeak • do n't squeak • them to squeak • did n't squeak • does n't squeak • enough to squeak

Figure 3: Example of a WORD search in COCA concerning the word squeak

- 4) Sharing and explaining (5-10'): The particular grammatical rules examined will be shared and confirmed by teachers, along with an explanation. It will also be an opportunity for students to debate and share their theories and hypotheses. For instance, students may notice that *-(i)on* is a noun-forming suffix, and that *-(i)* is associated with stems ending with *-s* (as in *discussion* or *division*) or *-t* (as in *competition* or *construction*).
- 5) Practice (10'): Students will engage in a game in which they will be able to practice word formation processes and use different language skills. The following game examples are suggested:
  - a) A matching card game for the affixation activity. Learners, at random, will be given cards with roots and affixes (Appendix 5). Each student will need to create as many words as possible, by asking others for cards.
  - b) News headlines for the derivation activity. Students select a card and they have to create two news headlines in two manners. First, including two uses of a given word, and secondly, substituting one of the terms with a synonym, as illustrated below in Figure 4:

“Outdoor dining boost in Spain as government extends COVID restrictions”

Use: governmental and boosting.

Example: “A governmental extension of COVID restrictions is boosting outdoor dining in Spain”

Substitute: restrictions

Example: “A governmental extension of COVID constraints is boosting outdoor dining in Spain”

Figure 4: Example of an accurately completed news headline activity

- c) New word entries for the onomatopoeias activity. Students will make up new word suggestions and create dictionary entries for them (e.g. *twimp*: 1. (noun) a sound produced by an object when it is introduced in a mass of water without splashing. 2. (verb) to produce a sound similar to that of an object when it is introduced in a mass of water without splashing).
- 6) Work on the portfolio (5'): Students will record the knowledge they have learnt, and reflect on their practice and attitude in a guided portfolio (Appendix 6).

#### 4.3.3 *Reviewing the contents (Session 5)*

A session could be devoted to practising the different word formation processes that have been studied beforehand. For this purpose, a session following the same pattern as that shown in Subsection 4.3.2 could be carried out. A text showing different word formation processes could be presented to the students, and, as a practical activity, they could take part in a game in which they would have to make their own hypotheses about words and whether they may exist or not in English. For instance, in this activity, learners could select two cards at random from a deck containing affixes and roots (previously printed out by the teacher) and combine them to create a word. They could write down the hypothetical word and decipher whether it exists or not, its meaning, and then prove their suggestions by searching the corpora.

## 5. CONCLUSION

This paper has attempted to shed light on how EFL teachers may introduce corpora in their classrooms. It has examined how corpora may be suitable tools to learn about word formation processes and to establish relations with other aspects of word depth for a better and more complete acquisition of lexis. More specifically, a particular learning proposal that sets the context for a group of Secondary Education students within the Spanish curricular framework has been presented.

The learning unit proposed has been developed in line with a combined method of DDL or inductive work and a more traditional, deductive methodology that is typically more familiar to students. In these sessions, focused on word formation, learning is facilitated through native English corpora. Thanks to these electronic collections, students may study words from multiple points of view, establishing links with diverse aspects of word knowledge, (e.g., meaning, pronunciation, collocates). Further, the resources and materials to be implemented allow placing vocabulary at the centre of the learning focus in a communicative manner. This way, learners explore new depths of knowledge about terms that they already know, while they learn new ones. These activities are based on a contextualised practice of word formation processes through the promotion of all language skills, in which a place for communicative situations has been granted. The contents would also potentially motivate students to elaborate hypotheses that require a substantial use of language. Finally, learners' ability to reflect on their learning process and product is put into practice with a personal portfolio. The aim is not only to introduce corpora, but to promote their use so that students continue consulting them autonomously during their life-long learning. The practical examples of the lexical aspect presented in this paper are just some of the many possible ways in which corpora may be introduced in the EFL classroom. In future stages, it would be interesting to implement the learning unit proposed in a real context, to monitor whether this methodology and the students' feedback is successful.

## REFERENCES

- Asención-Delaney, Y., Collentine, J. G., Collentine, K., Colmenares, J. & Plonsky, L. (2015). El potencial de la enseñanza del vocabulario basada en corpus: optimismo con precaución. *Journal of Spanish Language Teaching*, 2(2), 140-151. doi: 10.1080/23247797.2015.1105516
- Bacon, L. & Krpan, D. (2018). (Not) eating for the environment: The impact of restaurant menu design on vegetarian food choice. *Appetite*, 125, 190-200. doi: 10.1016/j.appet.2018.02.006
- Bogaards, P. (2001). Lexical units and the learning of foreign language vocabulary. *Studies in Second Language Acquisition*, 23, 321-343. doi: 10.1017/S0272263101003011
- Bolinger, C. (2020). Glen Carbon mayor displeased with trash hauler's pick-up postponement. *The Edwardsville Intelligencer*. Retrieved from <https://www.theintelligencer.com/news/article/Glen-Carbon-mayor-displeased-with-trash-15224273.php>
- Breyer, Y. (2009). Learning and teaching with corpora: reflections by student teachers. *Computer Assisted Language Learning*, 22(2), 153-172. doi: 10.1080/09588220902778328
- Canale, M. & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1(1), 1-47. doi: 10.1093/applin/I.1.1
- Caro, K. & Mendinueta, N. R. (2017). Lexis, lexical competence and lexical knowledge: a review. *Journal of Language Teaching and Research*, 8(2), 205-213. doi: 10.17507/jltr.0802.01
- Cofield, C. (2020). Where Are Stars Made? NASA's Spitzer Spies a Hot Spot. *NASA*. Retrieved from <https://www.nasa.gov/feature/jpl/where-are-stars-made-nasas-spitzer-spies-a-hot-spot>
- Conrad, S. & Levelle, K. (2008). In B. Spolsky & F. Hult (Eds.), *The Handbook of Educational Linguistics* (pp. 539-556). Oxford: Blackwell Publishing.
- Council of Europe. (2001). *Common European Framework of Reference for Languages: learning, teaching, assessment*. Cambridge: Cambridge University Press.
- Davies, M. (2004). British National Corpus (from Oxford University Press). Retrieved from <https://www.english-corpora.org/bnc/>
- Davies, M. (2008-). The Corpus of Contemporary American English (COCA). Retrieved from <https://www.english-corpora.org/coca/>
- Dewaele, J. (2009). Individual differences in second language acquisition. In W. C. Ritchie & T.K. Bhatia (Eds.), *The New Handbook of Second Language Acquisition* (pp. 623-646). Bingley: Emerald.
- Elgort, I. (2018). Teaching/developing vocabulary using ICTs and digital resources. *The TESOL Encyclopedia of English Language Teaching*, 1-15. doi: 10.1002/9781118784235.eelt0735
- Ellis, N. C. (2015). Implicit and explicit language learning: Their dynamic interface and complexity. In P. Rebuschat (Ed.), *Implicit and Explicit Learning of Languages* (pp. 1-24). Amsterdam: John Benjamins.



- Gabrielatos, C. (2005). Corpora and language teaching: Just a fling or wedding bells? *The Electronic Journal for English as a Second Language TESL-EJ*, 8(4). Retrieved from <http://tesl-ej.org/ej32/a1.html>
- Gaskell, D. & Cobb, T. (2004). Can learners use concordancer feedback for writing errors? *System*, 32, 301–319. doi: 10.1016/j.system.2004.04.001
- Gilquin, G. & Granger, S. (2010). How can data-driven learning be used in language teaching? In M. McCarthy & A. O’Keeffe (Eds.), *The Routledge handbook of corpus linguistics* (pp. 359-370). London: Routledge.
- Jamrus, M. H. M. & Razali, A. B. (2019). Using self-assessment as a tool for English language learning. *English Language Teaching*, 12(11), 64-73. doi: 10.5539/elt.v12n11p64
- Johns, T. (1991). Should you be persuaded: Two samples of data-driven learning materials. *English Language Research*, 4, 1-16.
- Kohonen, V. (2000). Student reflection in portfolio assessment: making language learning more visible. *Babylonia*, 1, 13-16. doi: 10.20533/licej.2040.2589.2018.0393
- Langley, L. (2015). Are lizards as silent as they seem? *National Geographic*. Retrieved from <https://www.nationalgeographic.com/animals/article/151024-animal-behavior-lizards-reptiles-geckos-science-anatomy>
- Laufer, B. (2005). Focus on form in second language vocabulary learning. *Eurosla yearbook*, 5(1), 223-250. doi: 10.1075/eurosla.5.11lau
- Lee, P. & Lin, H. (2019). The effect of the inductive and deductive data-driven learning (DDL) on vocabulary acquisition and retention. *System*, 81, 14-25. doi: 10.1016/j.system.2018.12.011
- Leńko-Szymańska, A. & Boulton, A. (Eds.). (2015). *Multiple Affordances of Language Corpora for Data-driven Learning* (Vol. 69). Amsterdam: John Benjamins.
- Lin, K. (2000). Chinese food cultural profile. *EthnoMed*. Retrieved from <https://ethnomed.org/resource/chinese-food-cultural-profile/>
- Ma, Q. & Kelly, P. (2006). Computer assisted vocabulary learning: Design and evaluation. *Computer Assisted Language Learning*, 19(1), 15-45. doi: 10.1080/09588220600803998
- Ma'arif, A. S., Abdullah, F., Fatimah, A. S. & Hidayati, A. N. (2021). Portfolio-based assessment in English language learning: Highlighting the students' perceptions. *J-SHMIC: Journal of English for Academic*, 8(1), 1-11. doi: 10.25299/jshmic.2021.vol8(1).6327
- Matos, S. C. G. (2013). *The use of corpora to teach English as a foreign language in Secondary Education in Portugal* (doctoral dissertation). Universidade Nova de Lisboa, Portugal.
- McCarthy, M. (2008). Accessing and interpreting corpus information in the teacher education context. *Language Teaching*, 41(4), 563-574. doi: 10.1017/S0261444808005247
- Meunier, F. (2011). Corpus linguistics and second/foreign language learning: exploring multiple paths. *Revista brasileira de linguística aplicada*, 11(2), 459-477. doi: 10.1590/S1984-63982011000200008
- Nation, I. S. P. (2000). *Learning Vocabulary in another Language*. Cambridge: Cambridge University Press.

- Oxford, R. (2003). *Language Learning Styles and Strategies: An Overview*. Oxford: Gala.
- Pérez-Paredes, P., Sánchez-Tornel, M., Alcaraz Calero, J. M. & Jiménez, P. A. (2011). Tracking learners' actual uses of corpora: guided vs non-guided corpus consultation. *Computer Assisted Language Learning*, 24(3), 233-253. doi: 10.1080/09588221.2010.539978
- Poole, R. (2018). *A Guide to Using Corpora for English Language Learners*. Edinburgh: Edinburgh University Press.
- Royal Decree 1105/2014 of December 26<sup>th</sup>, whereby the basic curricula of Secondary Education and Baccalaureate are established. (2014). Ministerio de Educación, Cultura y Deporte. «BOE» núm.3, de 3 de enero de 2015.
- Roca Varela, M. L. (2012). Corpus linguistics and language teaching: Learning English vocabulary through corpus work. *ES: Revista de filología inglesa*, 33, 285-300.
- Römer, U. (2006). Pedagogical applications of corpora: Some reflections on the current scope and a wish list for future developments. *Zeitschrift für Anglistik und Amerikanistik*, 54(2), 121-134. doi: 10.1515/zaa-2006-0204
- Römer, U. (2011). Corpus research applications in second language teaching. *Annual Review of Applied Linguistics*, 31, 205-225. doi: 10.24191/ijmal.v4i2.9449
- San Mateo-Valdehíta, A. (2013). El efecto de tres actividades centradas en las formas (focus on forms, fonfs): la selección de definiciones, la selección de ejemplos y la escritura de oraciones en el aprendizaje de vocabulario en segundas lenguas. *RAEL: Revista Electrónica de Lingüística Aplicada*, 12, 17-36.
- Selinker, L. (1972). Interlanguage. *IRAL: International Review of Applied Linguistics in Language Teaching*, 10, 1-4.
- Sinclair, J. (1999). The computer, the corpus and the theory of language. *Transiti Letterari e Culturali*, 2, 1-15.
- Tahaineh, Y. (2012). The awareness of the English word-formation mechanisms is a necessity to make an autonomous L2 learner in EFL context. *Journal of Language Teaching & Research*, 3(6), 1106- 1113. doi: 10.4304/jltr.3.6.1105-1113
- Vasu, K. A. P., Mei Fung, Y., Nimehchisalem, V. & Rashid, S. (2020). Self-regulated learning development in undergraduate ESL writing classrooms: Teacher feedback versus self-assessment. *RELC Journal*, 1-15. doi: 10.1177/0033688220957782
- Willis, D. (2003). *Rules, Patterns and Words: Grammar and Lexis in English Language Teaching*. Cambridge: Cambridge University Press.

## APPENDIX 1

### Chinese Food Cultural Profile

*"Fashion is in Europe, living is in America, but eating is in China."*

The phrase is a testament to the popularity of Chinese food around the world. Food is an important part of daily life for Chinese people. Chinese not only enjoy eating but believe eating good food can bring harmony and closeness to the family and relationships.

Shopping daily for fresh food is essential for all Chinese cooking. Unlike the fast food society of the U.S., the Chinese select live seafood, fresh meats and seasonal fruits and vegetables from the local market to ensure freshness. This means swimming fish, snappy crabs, and squawking chickens. Even prepared foods such as dim sum or BBQ duck for to go orders must gleam, glisten, and steam as if just taken out of the oven.

*Image 1. Fragment adapted from Lin (2000)*

### (Not) eating for the environment: the impact of restaurant menu design on vegetarian food choice

Meat is culturally accepted not only because it is important for social relationships, but also because it is universally regarded as a symbol of affluence and success (Smil, 2002). Indeed, the amount of meat consumed has been shown to rise with per capita income and has increased globally with GDP over the last 50 years (Tilman & Clark, 2014). Growth in meat consumption has been particularly rapid in some Northeast and Southeast Asian countries (e.g. China, Japan, Vietnam, and Thailand) as a result of economic development and globalization of the food industry (Nam, Jo, & Lee, 2010). In addition to cultural factors, lack of competence can also be an important barrier to reducing the intake of meat and eating more fruits and vegetables.

*Image 2. Fragment adapted from Bacon & Krpan (2018)*

## APPENDIX 2

- ❶ Function: LIST. Search for the term "library" in the BNC. Now search for the same word in the COCA. Would you say that "library" is frequent in each variant? In which one is it more frequent?
- ❷ Function: CHART. Search for the term "digital" in the BNC. In which genre is the term used more often? In which subject of non-academic publications is it mostly used?
- ❸ Function: WORD (COCA). You are reading a text and you come across the word "apron". Find out what it means by searching for this term. How many meanings does it have? How is it pronounced? Are there any synonyms for this word? Can you name some collocates? Topics? Clusters?
- ❹ Function: COLLOCATES. Think about the term "brilliant". Can you suggest other words that collocate with this term? Search for one collocation to the right of "brilliant" and see if they match. Now think about the word "job". Can you suggest any collocations? Search for one collocate to the left of the word. Do they match?
- ❺ Function: COMPARE- Do you think the words "attractive" and "beautiful" are synonyms? Why/why not? Search for these terms in BNC or COCA. What differences in meaning can you find based on their collocates?

*Image 3. Activity sequence adapted from Poole (2018)*

## APPENDIX 3

### Where Are Stars Made? NASA's Spitzer Spies a Hot Spot

Located about 17,000 light-years from Earth, in the direction of the constellation Aquila in the night sky, W51 is about 350 light-years – or about 2 quadrillion miles – across. It is almost invisible to telescopes that collect visible light (the kind human eyes detect), because that light is blocked by interstellar dust clouds that lie between W51 and Earth. But longer wavelengths of light, including radio and infrared, can pass unencumbered through the dust. When viewed in infrared by Spitzer, W51 is a spectacular sight: Its total infrared emission is the equivalent of 20 million Suns.

If you could see it with your naked eye, this dense cloud of gas and dust would appear about as large as the full Moon. The Orion Nebula – another well-known star-forming region and a favorite observing target for amateur astronomers – occupies about the same size area in the sky.

*Image 4. Affixation text adapted from Cofield (2020)*

### Glen Carbon mayor displeased with trash hauler's pick-up postponement

"These temporary changes will help keep our routes running smoothly and our communities clean. Customers are being notified of service changes by their municipality and/or directly from Republic Services.

"We recognize the coronavirus (COVID-19) situation is evolving rapidly and we are continuously evaluating the situation to ensure we can help keep our employees healthy and safe while working hard to ensure the highest quality customer service and minimal service disruption.

"Our top priority is the health and safety of our employees, and that is at the forefront of every decision we make. We are confident that we have the necessary plans and protocols in place for protecting employees from COVID-19."

*Image 5. Derivation text adapted from Bolinger (2020)*

### Are Lizards as Silent as They Seem?

"While it is true that most lizards are mute, many make sounds of various kinds," Robert Espinoza, a biologist at California State University, Northridge, explains via email. Geckos are the gabbiest, and some produce "a variety of chirps, clicks, and squeaks, some inaudible to humans," Espinoza says.

"The chirping, sometimes called 'barking,' of geckos is either a territorial or courtship display," to ward off other males or attract females, Peter Zani, a biologist at the University of Wisconsin–Stevens Point, says via email.

Some noteworthy noisemakers, Zani says, are Mediterranean house geckos, which squeak during fights and flirtatiously click to draw females. The turnip-tailed gecko of Central and South America makes territory-marking clicks thought to mimic insects. And the New Caledonian gecko, the largest gecko at 14 inches (36 centimeters), has a growl that earned it the local nickname of "the devil in the trees."

*Image 6. Onomatopoeias text adapted from Langley (2015)*

APPENDIX 4

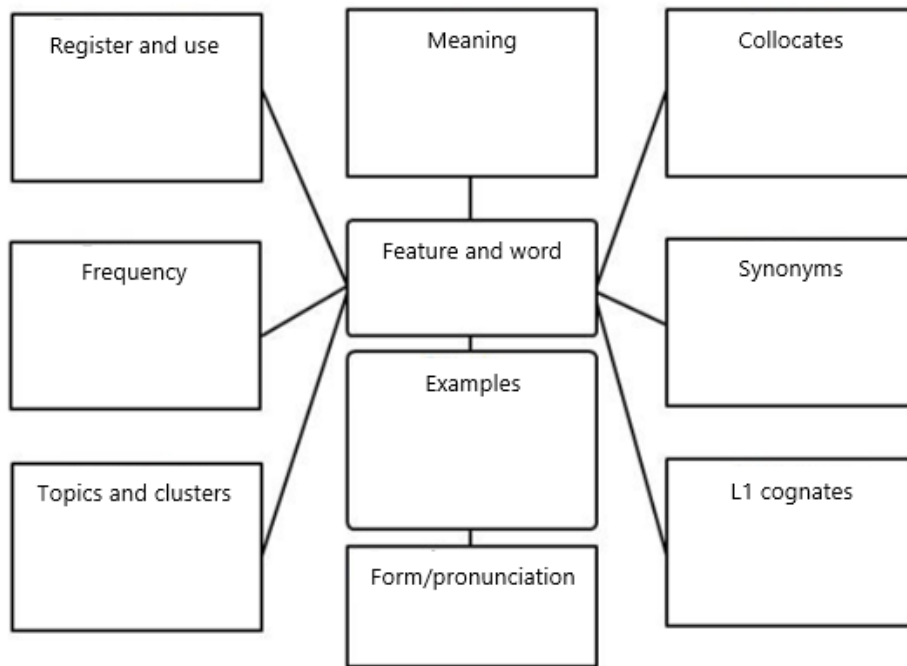


Image 7. Mind map sample

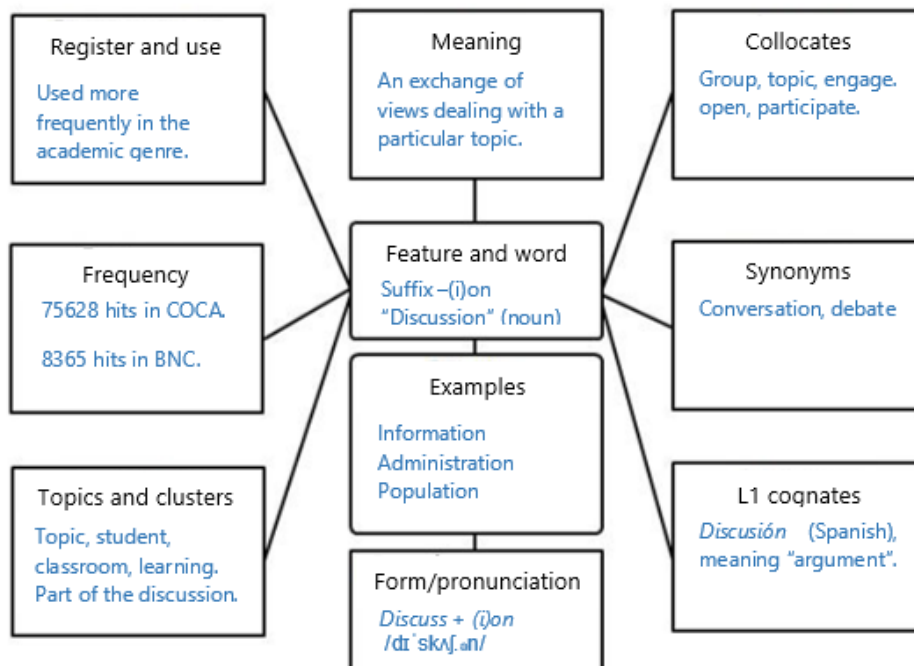


Image 8. Example of an accurately completed mind map

## APPENDIX 5

un-	safe
dis-	cheer
-less	job
-ful	tidy
-ment	like
-ly	home
	colour
	enjoy

*Image 9. Matching card game sample*

## APPENDIX 6

Session:
Achievements:
Rate the difficulty of this session:
Comments on your attitude in class:
I feel confident with:
I don't feel so confident with:
I may need help with:
Notes:

*Image 10. Daily portfolio rubric sample*