



**TRABAJO FINAL DE MÁSTER
MÁSTER EN ENSEÑANZA BILINGÜE PARA
PRIMARIA Y SECUNDARIA
FACULTAD DE EDUCACIÓN**



**RECEPTIVE VOCABULARY IN PRIMARY CLIL LEARNERS: CORRELATIONS
WITH L1 AND ACADEMIC ACHIEVEMENT.**

María Teresa Mendiola Romero
Prof. Dr. Rafael Alejo González
Departamento de Filología Inglesa
Máster en Enseñanza Bilingüe para Primaria
y Secundaria

CURSO 2019/2020
BADAJOZ
Convocatoria: junio 2020

Table of contents

1. Abstract	3
2. Resumen	4
3. Introduction	5
4. Justification	6
5. Aims of the study	7
6. Research questions	7
7. Theoretical Framework	8
7.1. CLIL in Second Language Acquisition	9
7.2. Vocabulary in Second Language Acquisition	12
7.2.1. Components of L2 vocabulary	12
7.2.2. Lexical competence	14
7.3.3. Comparing L1 and L2 vocabulary	19
7.3.4. Checking on lexical competence: vocabulary testing.....	22
8. The study	25
8.1. Participants	25
8.2. Instruments	26
8.3. Data collection.....	27
8.4. Results and discussion	28
8.4.1. RQ 1.....	30
8.4.2. RQ 2.....	32
8.4.3. RQ 3.....	34
9. Conclusions	38
10. References	41
11. Appendices	45

Figure 1: Distribution of students according to their scores in the tests	30
Figure 2: Test of normality English Test	45
Figure 3: Test of normality Spanish Test	45
Figure 4: Test of normality English Mark	46
Figure 5: Test of normality Natural Science Mark	46
Figure 6: Test of normality Social Science Mark	47
Figure 7: Test of normality Spanish Mark	47
Table 1: What is involved in knowing a word?	15
Table 2: Participants' information	25
Table 3: Descriptive Statistics	28
Table 4: Correlations	29
Table 5: Correlation between students' vocabulary and their marks (summary)	33
Table 6: Distribution of words in English	50
Table 7: Distribution of words in each set	51
Table 8: Percentages of correct answers in English and Spanish	53

1. Abstract

Vocabulary has been traditionally a neglected area in Second Language Acquisition (SLA). However, it is now clear for researchers and experts in the field that language is more than grammar. A new perspective emerged which understands language as grammaticalized lexis, emphasizing the importance of vocabulary acquisition. The perception of language has also influenced second language teaching, where new approaches like CLIL (Content and Language Integrated Learning) have arisen due to the need of learning foreign languages in our society. Research studies on CLIL have determined the advantages that this approach offers when it comes to vocabulary knowledge, although there are some opinions criticizing the effects on CLIL on L1 skills. In this study, *Peabody Picture Vocabulary Test (PPVT-III)* is used to compare CLIL primary school students' receptive vocabulary in an L1 and L2 (3rd grade). In addition, correlation among students' vocabulary and their marks is analysed. Finally, a qualitative study about the words that most student know and do not know is carried out. The results of the study show that there is no correlation between students' L1 and L2 vocabulary, but a positive correlation was found among students' L2 vocabulary and their marks. Also, a core and a peripheral vocabulary was identified. After analysing the results, it can be concluded that when studying vocabulary, it is important to consider individual differences and environmental influences. Moreover, the positive impact of L2 vocabulary in students' mark should be taken into account, which can be a great tool for teachers to empower their students. Finally, the qualitative analysis of words highlights the importance of cognates in vocabulary acquisition. At the end, some implications for teaching and the limitations of the present study are specified.

Keywords: SLA, receptive vocabulary, CLIL.

2. Resumen

El vocabulario ha sido tradicionalmente un área abandonada en la adquisición de segundas lenguas. Sin embargo, investigadores y expertos en este ámbito tienen claro ahora que el lenguaje es más que gramática. Una nueva perspectiva surgió entendiendo el lenguaje como léxico gramaticalizado, enfatizando la importancia de la adquisición de vocabulario. La percepción del lenguaje ha influido también en la enseñanza de segundas lenguas, donde nuevos enfoques como AICLE (Aprendizaje Integrado de Contenidos y Lenguas Extranjeras) han surgido de la necesidad de aprender lenguas extranjeras en la sociedad. Estudios de investigación en AICLE han determinado las ventajas que este enfoque ofrece en cuanto al conocimiento de vocabulario, aunque hay algunas opiniones criticando los efectos de AICLE en las habilidades de la L1. En este estudio, se utiliza *Peabody Test de Vocabulario en Imágenes (PPVT-III)* para comparar el vocabulario receptivo en la L1 y la L2 de estudiantes AICLE de educación primaria (3º). Además, se analiza la correlación entre el vocabulario y las notas de los estudiantes. Finalmente, se lleva a cabo un estudio cualitativo sobre las palabras que los estudiantes saben y aquellas que no saben. Los resultados del estudio muestran que no hay correlación entre el vocabulario de los estudiantes en la L1 y en la L2, pero se encuentra una correlación positiva entre el vocabulario de la L2 y las notas de los estudiantes. Además, se identifica el vocabulario común a todos los alumnos y el vocabulario periférico, esto es aquel vocabulario que sólo parte de ellos conoce. Una vez analizados los resultados, se puede concluir que es importante considerar las diferencias individuales e influencias del entorno al estudiar el vocabulario. Por otra parte, el impacto positivo del vocabulario de la L2 en las notas debería tenerse en cuenta pues, puede ser una gran herramienta para los profesores para empoderar a sus estudiantes. Finalmente, al analizar las palabras cualitativamente, se resalta la importancia de los cognados en la adquisición de vocabulario. Al final del documento, se especifican las implicaciones didácticas y limitaciones de este trabajo.

Palabras clave: adquisición de segundas lenguas, vocabulario receptivo, AICLE.

3. Introduction

Second Language Acquisition (SLA) has been a field of interest for decades but, the scientific discipline itself started in the mid 1900 with teachers, psychologists and linguists researching on different theories. Several second language teaching methods have emerged throughout the years according to research findings and practices on how to teach a second language (Cook, 2008). The grammar-translation method became very popular when SLA began and, it was used for decades. Translations from the L2 to the L1 was the most important aspect, ignoring other important issues in the language e.g. vocabulary or communication. After analysing its drawbacks, another method known as the Communicative Approach emphasised the need for real communication in a foreign language. It is from this point that vocabulary started to be considered an important aspect in SLA and researchers became more interested in the field (Castellano-Risco, Alejo-González & Piquer-Píriz, 2020).

The importance of vocabulary on foreign language proficiency is now clear for SLA experts. Research on vocabulary acquisition has influenced second language teachers and learners who are now emphasising vocabulary learning. However, there is still a lack of conclusive evidence about the most effective way to learn L2 vocabulary in an efficient way, which suggests that there is a need for further research on the field (Webb & Nation, 2017). As research evolves, new approaches to SLA have arisen. A good example of this is Content and Language Integrated Learning (CLIL), an educational approach that promotes learning content subject through an additional language. CLIL is widespread through Europe and there is ample research showing its positive impact on second language learning. One of its main strengths is related to vocabulary, as it has been confirmed through several studies that CLIL students present higher rates in receptive vocabulary knowledge than English as a Foreign Language (EFL) learners (e.g., Agustín-Llach & Canga Alonso, 2016; Jiménez Catalán & Ruiz de Zarobe, 2009). The main reason for this finding could be related to a richer language input received in CLIL practices or the amount of exposure time to the target language, although other aspects can also affect the results, such as the educational context (Castellano-Risco et al., 2020).

4. Justification

There are some studies supporting the fact that CLIL practices have a positive impact on learner's vocabulary knowledge (Castellano-Risco, 2018). However, there are other studies with less encouraging results about CLIL success (Bruton, 2011). Among the academic experts (teachers, linguists, and pedagogues), there are also mixed opinions. While some of them accept CLIL as an advantageous approach for learners' training, others are reticent to accept the different benefits that CLIL can offer as an educational approach. This situation is also noticeable in primary education where very young students (6 to 12 years old) learn different content subjects through an additional language.

Studies on CLIL learners in primary education are not as frequent as those with older learners mostly due to their limited proficiency level in a foreign language as well as difficulties with methodological aspects to test very young learners. In addition, there is a general belief among the educational community regarding the impact on students' L1 which many see as the major drawback of a CLIL approach. The need to produce empirical evidence of the results of CLIL in primary education still remains. Against this backdrop, this study modestly aims to test the receptive vocabulary knowledge of CLIL primary students in both their L1 and their L2, which will provide us with a great opportunity to establish whether there is a connection between both vocabulary levels and at the same time check if they relate to students' grades in the different subjects. In addition, there is also an analysis about which words CLIL learners know, and which one they are ignorant of, in order to have a qualitative view of CLIL learners' vocabulary knowledge in primary education.

This MA dissertation presents a study of receptive vocabulary knowledge in CLIL primary students (3rd grade) both in L1 and L2. First, it offers some key concepts about CLIL and its relation to vocabulary, followed by a review on the importance of vocabulary knowledge in SLA (components of L2 vocabulary and lexical competence) and complemented with a comparison between L1 and L2 vocabulary knowledge and vocabulary testing. Then, the data obtained from the two tests (one in English and one in Spanish) conducted with third grade students of primary education in a bilingual school in Extremadura is introduced. The data is then analysed comparing L1 and L2 vocabulary knowledge as well as their relationship with students' grades in different subjects and outlining the words that learners know and do not know. Finally, some conclusions about students' vocabulary knowledge are offered.

5. Aims of the study

The main aims in this study are:

1. To compare CLIL learners' receptive vocabulary in L1 and L2 in 3rd grade of primary education.
2. To relate receptive vocabulary in L1 to student's marks in subjects in Spanish and receptive vocabulary in L2 to student's marks in subjects in English.
3. To analyse words that most students know and those they are ignorant of to establish their different features.

6. Research questions

RQ1: Do children in 3rd grade of primary education with a high score of receptive vocabulary in their L1 also have a high score of receptive vocabulary in their L2?

RQ2: Are the scores in L1 vocabulary related to students' marks in the content subjects taught in Spanish? Are scores in L2 vocabulary related to students' marks in the content subjects taught in English?

RQ3: Is it possible to identify a core L2 vocabulary for 3rd graders, i.e. a group of English words known by most students at this educational stage versus a peripheral vocabulary, i.e. a group of words only known by some students? What are the different features that each of these two groups of words share?

7. Theoretical Framework

Language can be understood, according to a definition provided by the MacMillan Dictionary as “the method of human communication using spoken or written words” (‘Language’, n.d.). From this definition, the word human can be highlighted as a key factor when studying language. As Cook (2008) states, “language is at the centre of human life” (p. 1), and it can be considered the most unique characteristic of human beings differentiating them from other species. Some people not only learn a first language but also develop a competence in several languages. SLA understood as “the process by which people learn a language that is not their native language” (‘Second Language Acquisition’, n.d.) is increasing in popularity around the world. As Cook (2008) clearly explains:

In a world where probably more people speak two languages than one, the acquisition and use of second languages are vital to the everyday lives of millions; monolinguals are becoming almost an endangered species. Helping people acquire second languages more effectively is an important task for the twenty-first century. (Cook, 2008, p. 1)

In this task of supporting people in their SLA process, one of the most suitable tools is education. Thus, it is not surprising that some educational approaches such as CLIL have emerged as a possible solution to the current need of learning foreign languages.

Learning a language both as an L1 or as an L2 involves developing different receptive and productive skills (listening, reading, speaking, and writing) as well as other aspects involved in the communicative competence (discourse, strategic, socio-linguistic and grammatical competence). In previous decades, great importance was given to grammar and that was reflected in the teaching methods used. However, some experts in the field affirm that there is one key aspect, particularly important in SLA, i.e. vocabulary (Nation, 2001). Words are essential in communication and it is through vocabulary that content learning can be ensured. This is why there are many challenges that learners will face when learning vocabulary (Lightbown & Spada, 2013; Webb & Nation, 2017).

It is now clear that grammar itself is not the most important aspect of language which is enhancing the importance of vocabulary. As Lewis (1993) states, “language consists of grammaticalized lexis, not lexical grammar” (p.34). This means that we learn vocabulary associated to some structures and not structures in which vocabulary is inserted.

The importance of learning vocabulary both in an L1 and L2 is widely accepted and there are now many researchers interested in this field (e.g., Lightbown & Spada, 2013). However, in our society learning only one language is not enough anymore. Learning an L2 is a real need

in the actual society, for this reason, some educational approaches like CLIL have arisen to face this situation obtaining encouraging results, especially in relation to vocabulary acquisition.

7.1. CLIL in Second Language Acquisition

Content and Language Integrated Learning, or CLIL, is a term originally coined in Europe in 1994. This practice has existed in different cultures over time but the growing interest in this approach has led to the recent name recognition (Mehisto, Marsh, & Frigols, 2008). However, the definitions of CLIL are still varied and internally ambiguous, which has prevented experts from establishing general characteristics that are unique to this approach (Cenoz, Genesee, & Gorter, 2014). The most widely used definition nowadays considers CLIL as a “dual-focused educational approach in which an additional language is used for the learning and teaching of both content and language” (Coyle, Hood, & Marsh, 2010, p. 1). This definition focuses on content and language with the same emphasis. Other authors agree on considering CLIL as an approach that gives special attention to these two aspects, but also include a third element, learning skills. This is known as the CLIL triad that supports content and language goals. Thus, CLIL is defined as an umbrella term that covers varied educational approaches (Mehisto et al., 2008).

The interpretations of CLIL as an educational approach are the most common throughout the experts in the field. Some authors understand CLIL as a holistic approach, emerging from interdisciplinary research, adapted to new forms of knowledge (Coyle et al., 2010). CLIL developed as a response to European needs distinctive of the Knowledge Society, based on integrated learning demands from globalization and mobility. CLIL is also designed to respond to cultural demands (the need of second language proficiency), as well as the increasing use of ICTs and the need of immediacy of purpose when learning (Coyle et al., 2010; Mehisto et al., 2008). This integration is what makes CLIL unique as an educational approach. According to Coyle et al. (2010), the most relevant characteristics are related to factors such as a great increase in linguistic competence, developing cognitive flexibility, encouraging motivation to learn a language based on authenticity and acquiring a language in real life situations, close to natural environments and distant from traditional instructional settings. Other authors like Mehisto et al. (2008), agree with the previous ideas and add that apart from academic content goals, language proficiency, and cognitive and social skills and habits, CLIL emphasises L1 competence, a goal not followed by language teaching.

There are numerous books, articles, and studies in the academic and pedagogical field about the theoretical basis of CLIL and many of them report the effectiveness of CLIL settings in relation to the development of learners' communicative competence. However, less positive results have been obtained in relation to productive skills. Most of these studies had been conducted in North American bilingual education contexts, which implies that results cannot be transferred to others settings like Europe (Pérez Cañado, 2012). Although there is a great need of empirical research in Europe, some studies have focused on observing teachers and students in a CLIL context while others compare student's mastery of the second language in EFL and CLIL contexts. In the latter, similar results to the ones in North America were obtained when comparing oral second language competence. CLIL students usually outperform non-CLIL students (Jiménez Catalán & Ruiz de Zarobe, 2009).

There are also some studies assessing vocabulary knowledge in CLIL settings. One of them was carried out in Spain by Jiménez Catalán & Ruiz de Zarobe (2009) who compared receptive vocabulary in EFL and CLIL learners. After analysing the results, they claim that "CLIL instruction is more effective than ES (English as a subject, non-CLIL) instruction as it is the CLIL group that achieves better results in both tests" (p. 87), a conclusion that Agustín-Llach & Canga Alonso (2016) were also able to reach in their study. The same results were found in a study carried out in Finland where a comparison between students' vocabulary knowledge in CLIL and mainstream education was completed. In this study, CLIL learners scored higher than their peers who did not attend CLIL education. Another interesting conclusion from this article is the fact that, as expected, CLIL learners "lag behind native speaker word knowledge" (Merikivi & Pietilä, 2014, p. 495). These encouraging findings supporting CLIL instruction are not only related to the educational tools used in CLIL but also to the benefits of longer exposure time to the target language and due to the meaningful learning context that CLIL offers (Canga Alonso, 2015).

CLIL is becoming very popular in Spain, which, encouraged by education authorities promoting multilingual policies, is becoming a leader in Europe when it comes to CLIL practice and research. The different regions in Spain have different characteristics with significant differences between bilingual and monolingual areas. Traditionally, bilingual areas have implemented and conducted research to a great extent, but in monolingual communities, even though CLIL tradition is more recent, research has been thriving (Pérez Cañado, 2012).

Focusing on the specific geographical context of the present study, the characteristics of CLIL in Extremadura as a monolingual community have also been discussed by some authors. CLIL was officially implemented in the academic year 2004-2005 with a project promoted by

education authorities in the region called bilingual section projects. Nevertheless, experience in teaching content subjects through a foreign language started in the academic year 1996-1997 in two schools in the region. From this date, encouraging second language learning was a priority in the region, expanding the teaching of a second language from nursery education to the whole educational system. In addition, not only one foreign language but two foreign languages were gradually introduced in the curriculum for primary and secondary education. According to Alejo & Piquer-Píriz (2010), some of the main characteristics of CLIL in Extremadura are: 1) A minimum of 2 and a maximum of 3 CLIL subjects are taught through an additional language; 2) Students have at least one session a week in the L2; 3) CLIL students must take an additional foreign language; 4) CLIL teachers need language certification; and 5) In most cases, English is the vehicular language used although French and Portuguese CLIL sections are allowed.

In a research study conducted in Andalusia, Canary Islands and Extremadura –three monolingual communities– empirical evidence about the positive effects of CLIL programs in foreign language competence was obtained. In this longitudinal study, Pérez Cañado (2018), compared the results obtained by 1033 CLIL learners and 991 EFL students with a pre-test, post-test and delayed post-test. All variables in the study were controlled to ensure that only CLIL had an impact on the results. She finally concluded that there was a “linguistic competence differential between CLIL and EFL groups, in favor of the former” (p.67) and that CLIL and not any other variable was the sole responsible factor for these results. Regarding vocabulary, in a study carried out in Extremadura by Castellano-Risco (2018) where she compared receptive vocabulary size in CLIL and non-CLIL learners, a positive correlation between participating in CLIL and presenting a larger receptive vocabulary size was found. The author explains that the reasons for this finding could be the greater exposure time to the target language in school and the specific words used in the various subjects that they learned through English within the CLIL programme.

To conclude, it could be said that results obtained in Extremadura match those obtained in other regions of Spain and other countries in Europe regarding the benefits of CLIL in L2 competence and more specifically in vocabulary knowledge. As Pérez Cañado (2018) clearly states in her study, it is only with empirical data that the conditions for CLIL to have a positive effect on learning can be known.

7.2. Vocabulary in Second Language Acquisition

7.2.1. Components of L2 vocabulary

As mentioned above, learning vocabulary is crucial to develop L2 linguistic competence. Learners need words to communicate their own thoughts and opinions, but words are also needed in order to understand written and oral texts. According to Webb & Nation (2017), vocabulary knowledge can be the basis for other aspects of language such as phonology, syntax, morphology, or pragmatics. In school years, vocabulary takes a very important role in content learning; useful words are highlighted, words are introduced before the topic to ensure comprehension, teachers usually work on spelling, pronunciation and meaning and learners are particularly aware of the importance of learning vocabulary as a means of accessing content. Research has shown that the degree of vocabulary knowledge that students have in an L2 is correlated with their grades in the L2 (Laufer & Goldstein, 2004). However, learning vocabulary is not an easy task for learners specially in a formal context where exposure to the language is very limited, while it can be an encouraging activity for learners who can clearly see their progress on vocabulary learning in contrast to grammar learning where the progress is not always so straightforward. As Webb & Nation (2017) state “perhaps no aspect of language learning is as satisfying as vocabulary learning” (p. 6).

In language learning, knowing vocabulary is equated to knowing words. But, what does it actually mean to know a word? In order to analyse this question, first some terminology related to the field needs to be clarified. When doing research on vocabulary, it is important to know what counts as a word. There are different ways of counting words in a corpus:

- Tokens: “every word form in a spoken or written text and if the same word form occurs more than once, then each occurrence of it is counted” (Nation, 2001, p. 7).
- Types: all words in a text without counting words that are repeated (Nation, 2001).
- Lemmas: “...a headword and some of its inflected and reduced (n’t) forms.” (Nation, 2001, p. 7). For example: *walk, walks, walked* and *walking*.
- Word families: “...a headword, its inflected forms and its closely related derived forms” (Nation, 2001, p. 8). For example: *work, worker, rework, workshop, workmanship, working, works* and *worked*, among others.

According to Nation (2001), researchers were able to determine that there are around 114,000 word families in English, which is an overwhelming number of words for native

speakers and even more for second language learners. In fact, educated native speakers of English only know around 20,000 word families (Nation, 2001) and this is considerably more demanding to achieve for second language learners, if at all possible. It has been estimated that vocabulary grows at 1,000 word families per year in an L1, so by the age of 20, educated native speakers (university students) should have reached the maximum vocabulary size (Nation, 2001). However, for various reasons this is not the case for second language learners, where vocabulary growth takes place at a slower rate (Miralpeix, 2019).

The positive aspect is that speakers do not need to know all existing words to use and understand a first or foreign language. In order to determine which are the most valued words, researchers classify words according to their frequency in four different groups: high-frequency words, academic words, technical words and low-frequency words.

High-frequency words are very important because “they cover a very large proportion of the running words in spoken and written texts and occur in all kinds of uses of the language” (Nation, 2001, p. 13). Researchers have different opinions when defining the limit of high-frequency words. Nation (2001) considers as high-frequency words the most frequent 2,000 word families in English and their importance lies in that around 80% of the words in most texts belong to this frequency band. This means that 4 in every 5 words found will be known by a student with this vocabulary level but, at the same time this also means that this student will not know as many as 1 word in 5 of a running text. The 1,000 most frequent word families have a text coverage of 77% and the next 1,000 word families represent around 5% of the running words in academic texts. In other words, due to the high percentage that they represent in a text (Nation, 2001), it is crucial for teachers to make sure their learners have learned these frequency bands. Other researchers (e.g. Schmitt & Schmitt, 2014) consider the first 3,000 word families as high-frequency words, arguing that they represent 98% of most reading materials and 95% of spoken discourse. The importance of this is that researchers indicate that learners are likely to understand a text if they know these percentages of vocabulary. For this reason, “learning the 3,000 most frequent word families would therefore seem to have some very important pedagogical implications” (Webb & Nation, 2017, p. 12).

Academic and technical words have some similarities and differences. Both are considered specialized vocabulary, in order to increase high-frequency words for specific purposes. The difference is that while academic words are frequent in most academic texts in all disciplines and infrequent in non-academic texts, technical words are only those related to a specific topic of a discipline (Webb & Nation, 2017). Academic words can be a challenge for L1 and L2 learners because this type of language, the language of schooling, is fairly distant from

everyday language or the language spoken at home. The teacher has to act as a mediator in order to scaffold language from everyday language into academic language, especially for those students who cannot get support at home (Schleppegrell, 2006). In contrast, technical words are more noticeable for students, because they are unique of a field, and teachers should not put the focus on them as they will be learned through the study of the topic. Knowing a technical word involves having good command of the subject-matter. Therefore, teachers should not forget to teach these words either, as students' understanding of technical words is crucial for the understanding of specific content (Nation, 2001). This is particularly important in CLIL where students are learning content through an additional language.

Low-frequency words are “a very large group of words that occur very infrequently and cover only a small proportion of any text” (Nation, 2001, p. 19). Some low-frequency words are those with a medium frequency within a language. Other words with low frequency are proper names and the vast number of extremely low-frequency words rarely used within a language. But, on specific occasions, low-frequency words can also be considered technical words, depending on the field of specialization. As a result, teachers should not aim to teach all low-frequency words appearing in the texts they use but instead provide students with some strategies in order to fully understand the text. Some of these strategies would be guessing meaning from the context or using word parts, to help them keep improving their vocabulary knowledge (Nation, 2001).

Word frequency is crucial in CLIL where students are exposed to high-frequency words but also to technical words, depending on the subject they learn through an L2. Analysing word-frequency would be very useful for CLIL teachers to decide the most appropriate words to teach.

7.2.2. Lexical competence

Once we have analysed how the vocabulary of a language can be characterized by using frequency, it is important to establish how learners and users become acquainted with words. Namely, it is now important to have a look at the research strand that pays attention to how those words become part of the mental lexicon of an L2 speaker and fit in his or her overall linguistic competence.

Here we need to consider two concepts introduced by Anderson and Freebody (1981 cited by Li & Kirby, 2014, p.612) that should be known in relation to vocabulary knowledge: vocabulary breadth and vocabulary depth. Vocabulary breadth is a quantitative aspect that

refers to the number of words that a person knows while vocabulary depth is a qualitative aspect that refers to how well a person knows words. Most researchers agree on how to count vocabulary breadth by the different ways of measuring words (Li & Kirby, 2014). However, vocabulary depth is more complex to define and takes us back to the question of what it is to know a word.

According to Nation (2001), there are three basic elements involved in knowing a word: form, meaning and use. Form refers to the spoken sound, pronunciation, written appearance, spelling, and word parts. Besides, meaning relates to the specific meaning of a word in a particular context, knowing which word to use for expressing specific messages, the concepts and referents that the word can be related to and associations with other words (thinking of connected words or synonyms). Finally, using a word can be understood as knowing its grammatical functions, collocations and constraints of use. Vocabulary depth is related to the knowledge of all these aspects involved in knowing a word. The more aspects a learner knows about a word, the greater his/her vocabulary depth. Nation (2001) includes a very clarifying table (Table 1) where the different elements (form, meaning and use) are explained. He also specifies its different features and makes a very useful distinction between receptive and productive vocabulary.

Table 1

What is involved in knowing a word?

Form	Spoken	R	What does the word sound like?
		P	How is the word pronounced?
	Written	R	What does the word look like?
		P	How is the word written and spelled?
	Word parts	R	What parts are recognizable in this word?
		P	What word parts are needed to express meaning?
Meaning	Form and Meaning	R	What meaning does this word form signal?
		P	What word form can be used to express this meaning?
	Concept and References	R	What is included in the concept?
		P	What items can the concept refer to?
	Associations	R	What other words does this word make us think of?
		P	What other words could we use instead of this one?
Use	Grammatical functions	R	In what patterns does the word occur?
		P	In what patterns must we use this word?
	Collocations	R	What words or types of word occur with this one?
		P	What words or types of words must we use with this one?
	Constraints on use	R	Where, when and how often would we meet this word?
		P	Where, when and how often can we use this word?

Note. R = receptive, P = productive

Table 1: *What is involved in knowing a word?*

Source: Nation (2001, p.27)

The distinction between receptive and productive vocabulary is very useful in order to analyse vocabulary learning. These terms can be applied in different areas of language knowledge and use, and they are related to the concepts of receptive skills (listening and reading) and productive skills (speaking and writing). According to some authors (Palmer, 1921; West, 1938; Crow, 1986 cited by Nation, 2001, p. 24), receptive vocabulary refers to the words that learners receive through oral or written input and try to understand them. Productive vocabulary is related to words learners can produce orally or in written form in order to convey information to others. However, Nation (2001) claims that this terminology is not completely suitable because “there are productive features in the receptive skills” (p. 24). For example, listening and reading involve producing meaning (Nation, 2001). There are two main positions discussed by experts in the field. One view is explained by Melka Teichroew (1982 cited by Nation, 2001, p.25). She holds that the distinction receptive/productive is arbitrary in relation to what is involved in knowing a word and that it would be preferable to consider vocabulary as a continuum, i.e., as a scale of knowledge. Nevertheless, many authors are still in favour of the distinction receptive/ productive vocabulary (Nation, 2001).

Other authors like Corson (1995 cited by Nation 2001, p.25) prefer using the term passive to designate receptive vocabulary and active to designate productive vocabulary, so in a way they can be considered synonyms. Even so, some researchers do not agree with this terminology either as they do not relate listening and reading with all the characteristics of the word passive. In the same way, there are also different positions about this distinction that some authors hold. Meara (1990 cited by Nation, 2001, p.25) states that active vocabulary is the one activated by other words while passive vocabulary is the one activated by external input. However, Nation (2001) claims that external stimuli are difficult to control, so it would be very difficult to determine to what extent active vocabulary is influenced by external elements. In addition, Corson (1995 cited by Nation, 2001, p.25) gives more importance to the idea of use and holds another point of view. To exemplify this, he says that although having a good level of passive vocabulary if it is not used, it will never be active. In addition, he claims that productive vocabulary includes receptive vocabulary and three other types of vocabulary, i.e., words partly known, low-frequency words not ready to use and words that are avoided (Nation, 2001).

In short, following Pignot-Shahov (2012), we can conclude that although the distinction between receptive and productive vocabulary –also designated as active and passive vocabulary– can have different interpretations from different researchers (from the ones that see it as a continuum to the ones that see it as two different categories), they all accept that this

is a dimension to be considered in the field of vocabulary acquisition. Nevertheless, there is still no agreement in the definition of the terms.

Taking all these aspects into account, Ellis & Beaton (1993), who use the distinction between receptive and productive vocabulary as a useful tool to study vocabulary acquisition, reach the conclusion that receptive learning is easier than productive learning. They provide the different reasons that support their findings. First, productive learning is more difficult because it requires knowing more distinctive aspects of words than receptive vocabulary, where less precision is needed. Second, when using language, receptive use is practiced more frequently than productive use. Third, when learning a new word in a foreign language, receptive vocabulary only has a simple link to the L1 word (translation) whereas productive vocabulary has several associations with the L1 word (collocations, synonyms, opposites, etc.), which makes it more difficult. Finally, learners are motivated to use some words productively but not others. Even though those words are known, if they are never used, they cannot be considered as productive vocabulary.

There are many authors who share the idea that productive vocabulary is acquired more slowly than receptive vocabulary and, to prove this, several studies measuring receptive and productive vocabulary size have been carried out. A good example is a study conducted by Webb (2008), which used translation tests to measure productive and receptive knowledge of meaning and use. This research revealed that receptive vocabulary size was larger than productive vocabulary size and that learners with higher scores in receptive vocabulary had also higher scores in productive vocabulary size.

In another study carried out by Stoddard (1929 cited by Nation, 2001, p.31), one of the first researchers to compare receptive and productive vocabulary, similar results were obtained. He divided learners in two groups and instructed them on vocabulary. One group practiced receptive vocabulary and the other group practiced productive vocabulary. Then, both groups completed receptive and productive tests. Learners got higher scores in the type of vocabulary learning that they had practiced. However, considering global scores, learners scored higher in receptive tests. Thus, he concluded that receptive tests are easier than productive tests because the global scores were higher. Waring (1997 cited by Nation, 2001, p. 32) conducted a very similar study obtaining identical results (Nation, 2001). In the same way, a study conducted by Zhou (2010) comparing academic vocabulary size in Chinese EFL students, results were alike. Learners obtained higher results in the receptive vocabulary test and learners with higher scores in receptive vocabulary also scored higher in the productive vocabulary test.

To conclude, experts in the field agree that productive knowledge requires previous receptive vocabulary knowledge. Researchers have shown through different studies that a word is easier to remember receptively than productively. Thus, it is not surprising to find statements in the sense that receptive vocabulary is acquired more easily than productive vocabulary knowledge (Webb & Nation, 2017).

Once we have given an answer to the question of what is involved in knowing a word, a very important concept for teachers and learners needs to be discussed. There are many aspects involved in knowing a word but, how easy or difficult is it to learn a specific vocabulary item? This is what is referred by the concept of learning burden. “The ‘learning burden’ of a word is the amount of effort required to learn it” (Nation, 2001, p. 23). The learning burden of a word is not necessarily the same for all learners, i.e., not all words pose the same amount of difficulty. According to Webb & Nation (2017) there are different factors affecting the difficulty involved in learning a word. First, we have the similarity with the L1 of the learner. When learning an L2, learners already have an existing system of knowledge (the one from the L1). The learning burden of a word will be higher when similarities between the L1 and the L2 are few and the linguistic systems of both languages are different. Second, there is the question of words following rules and exceptions. When learning a word, it is possible to focus on two complementary levels: a first level which means focusing on single words and a second level which involves focusing on rules that can be applied to many words. Finally, other aspects related to form, meaning and use should also be taken into account. Any of the specific aspects involved in knowing a word explained in Table 1 can make a word difficult to learn.

However, among all the aspects mentioned, the most influential factor is probably the similarity between the first and second language of the learner. As stated by Nation (2001), “for learners whose first language is closely related to the second language, the learning burden of most words will be light” (p. 24), which means that the opposite is also true. Teachers should always consider the learning burden of words in relation to their learners’ needs and offer strategies to their students so that learning becomes easier. Some techniques to help students reduce the learning burden involve, for example, highlighting patterns and analogies within their L2 or specifying connections with their L1 (Nation, 2001).

It is worth emphasizing that when learning vocabulary in an L2, there are many elements that influence lexical competence. Learners need complete information about words to know how to use them (vocabulary breadth). As discussed above, knowing the meaning of a word might not always be enough. For this reason, when it comes to vocabulary learning, CLIL education can bring about benefits for learners as it affords different L2 learning contexts which

can help students gather more information about words and consequently acquire a better lexical competence.

7.3.3. Comparing L1 and L2 vocabulary

Vocabulary knowledge is crucial in an L1 and an L2 in order to ensure communication, although the process of learning vocabulary is not identical in native and non-native speakers. According to Miralpeix (2019), there are some similarities between learning words in an L1 and an L2. First, learning vocabulary is a difficult task in all languages and at all ages. Thousands of words exist in all languages and this comprises a challenge for learners who have to deal with form, meaning and use (Meara, 1988; Singleton, 1999 cited by Miralpeix, 2019, p.190). Second, there is a relationship between L1 and L2 vocabulary with respect to social class and parental education that affects both first and second language vocabulary achievement. That means that higher social class and parental education positively affects vocabulary knowledge both in a first and second language. Also, vocabulary size can predict foreign language skills 10 years later (Skehan, 1989 cited by Miralpeix, 2019, p.190). Finally, without considering individual differences, good learners at learning vocabulary in their own language are more successful at learning vocabulary in another language and, vice versa, learners with problems to acquire vocabulary in their L1 will also have difficulties in learning vocabulary in an L2 (Meara, 1996). However, in a research study conducted in France with pre-literate children in relation to vocabulary, it was found that good L2 learners were not those with a larger L1 lexicon. This highlights the importance of analysing individual differences and environmental influences (Leśniewska & Pichette, 2016).

There are also some differences between L1 and L2 vocabulary learning. To start with, the way vocabulary is learned in an L1 is different to the way it is learned in an L2 due to three main factors: the learner, the context and the type of language involved. L2 learners have already acquired some strategies for learning vocabulary in their L1. Besides, L2 learners are usually more mature when they start learning and they already have some lexical knowledge from their L1 (Meara, 1988 cited by Miralpeix, 2019, p.190). In addition, a key aspect is that L2 exposure time is usually more limited and learning just takes place in a formal setting. Furthermore, the type of language involved is not the same in a first and a second language. It is claimed that “children tend to learn L1 vocabulary incidentally, as massive oral and written input is readily available to them, although deliberate learning can also enhance the lexical

growth of L1 learners during their school years” (Miralpeix, 2019, p.190). On the other hand, L2 learners learning in a formal setting cannot get enough exposure time in order to allow incidental learning to happen (Waring & Nation, 2004 cited by Miralpeix, 2019, p.190). Thus, vocabulary learning in a second language usually takes place through deliberate learning. Finally, it is important to know what types of languages are involved because if the L1 and L2 are languages typologically close, some words will be less difficult to learn, especially cognates (Nation, 1990 cited by Miralpeix, 2019, p.190).

Once the main similarities and differences have been discussed, vocabulary size and growth in a first and second language can be examined. As it was mentioned in previous sections, vocabulary size in L1 English is known to be around 4,000-5,000 word families for 5 year old children assuming that children learn about 1,000 word families per year in their L1. When it comes to the second language, it is clear that vocabulary acquisition cannot take place at the same pace as in the first language. Experts in the field estimate that vocabulary grows at a rate of around 500 word families per year in an L2 in the adequate context and under the right conditions (Webb & Nation, 2017). Considering what Schmitt & Schmitt (2014) establish as the minimum words necessary for communication, 5 or 6 years of L2 instruction will be needed in order to learn 3,000 word families, even though this is a demanding goal for some learners that never manage to achieve it. The main reason for this big difference between L1 and L2 vocabulary is related once more to exposure time and the limitations encountered in formal contexts, e.g., the reduced amount of input available for learners (Miralpeix, 2019).

Some research has been carried out to examine vocabulary size and growth in natural settings, but results are very difficult to relate and compare to other findings due to the varied formats used in the studies. Different studies not only used different tests to assess vocabulary but also results were analysed according to different criteria which makes it difficult to make valid general claims. On the other hand, there are many studies about vocabulary size and growth in formal contexts. Miralpeix (2019) offers a very interesting list of studies measuring L2 vocabulary in formal settings where some of them were conducted in primary education. In Hungary, a study carried out by Orosz (2009 cited by Miralpeix, 2019, p.197) measuring receptive vocabulary in grade 3 showed that children knew around 500 words. A similar study carried out in Spain in grade 6 by Jiménez et al. (2006 cited by Miralpeix, 2019, p.197) estimated a vocabulary size lower than 1,000 words (Miralpeix, 2019).

In her work, Miralpeix (2019) concludes that there is a problem in measuring vocabulary size related to methodological aspects. First, there is a lack of studies on specific areas and

second, results found in different studies with similar characteristics show great differences, which makes it very difficult to offer a valid estimation of vocabulary size in specific settings.

Another important issue related to frequency in vocabulary knowledge is the fact that high-frequency words in the first language might not be high-frequency words in a second language, especially if the L2 is learned in formal contexts where exposure time is very limited. This means that children might be familiar with different words in the L1 and in the L2. As Goriot et al. (2018) clearly state:

Words that are frequent in English as an L1 are not necessarily frequent in English as an L2, but frequency measures for English as an L2 are not available. It is therefore difficult to determine to which specific words children are exposed, and with what frequency. (Goriot et al., 2018, p.3)

According to these authors there are two predicting factors in L2 word knowledge for children learning English as an L2. First, we have high-frequency words in the first language (specially in early stages), which then become less important in favour of high-frequency of the L2 words (in this case, English). Then, we can also consider the similarities between L1 and L2 words, particularly cognates, which are easily recognized by students. In fact, in a study conducted in Europe, it was found that children with a first language typologically close to English, i.e., a language with the same origin and common characteristics, performed better than those whose L1 was more distant (Lindgren & Muñoz, 2013).

The relationship between first and second language is especially relevant in CLIL contexts, where a vehicular language is used to teach content subject. In contrast to what usually happens in EFL, where students learn new words as a complement of an already existing concept and word in the first language, CLIL students “simultaneously acquire new subject knowledge (e.g., new concepts) and the language to communicate this knowledge” (Gablasova, 2014, p. 977). When learning subject specific content, children encounter many new technical words that they have to learn in order to understand the concepts. In a study conducted with CLIL and non-CLIL students, a difference between learning new technical words in a first and second language was found. First, it was discovered that students learning through a first language outperformed those learning through a second language in form-meaning association, which means that L2 learners were able to achieve a lower vocabulary breadth. In relation to vocabulary depth, there was also a difference between both groups. L2 learners showed a “more limited conceptual knowledge of the technical terms” (Gablasova, 2014, p. 986) and they were able to recall fewer meaning components. In addition, after a period of time CLIL learners had more difficulties to remember technical words than students learning through a first language (Gablasova, 2014).

Results obtained in this research study are not surprising, as the challenges of learning vocabulary have been discussed with a special focus on foreign language, which makes it even more demanding. Nevertheless, the most important aspect for CLIL teachers from this study is to be aware of the challenges their students face when learning through an additional language. Teachers have a key role in helping their students acquire the vocabulary they need to learn content using specific strategies and scaffolding techniques “concerned with fostering both the connections between the words and their meanings as well as in-depth understanding and expressible knowledge of words” (Gablasova, 2014, p.988).

From the results obtained in the different research studies, it can be concluded that vocabulary acquisition is a different process in a first and a second language. As CLIL students face the challenge of learning content through a non-native language, teachers should be prepared to help their students in the difficult task of learning vocabulary, otherwise learning and understanding cannot take place. If CLIL is properly implemented, students will be positively influenced by this approach regarding vocabulary acquisition which will have a favourable influence in second language proficiency.

7.3.4. Checking on lexical competence: vocabulary testing

In the previous sections we hope to have been able to establish vocabulary acquisition as a very important aspect in SLA. Thus, the main features involved in vocabulary learning have been discussed together with the main difficulties it poses. For students learning a second language in formal settings, the role of the teacher is crucial to help them acquire the vocabulary required for the specific contexts where they need to communicate. This leads us to another key question: how can teachers check their students’ vocabulary knowledge? Even though the answer might sound simple, it is more complex than it seems.

Teachers should become researchers in their own classroom to offer students the best quality education. Traditionally, students’ knowledge has only been assessed to obtain the required grades but, assessing knowledge with the purpose of helping students improve and at the same time help instructors adapt teaching practices can be very enriching. Testing vocabulary is a great method for second language teachers to be aware of what their students know and what they can understand.

According to Nation (2001), “testing vocabulary is similar to testing in other areas of language knowledge and use” (p.344). There are some criteria that need to be considered when

deciding a test or assessing vocabulary knowledge. For Read (2019), the most important aspects are deciding the purpose of assessment and choosing the most suitable format.

Not all vocabulary tests measure the same dimensions of vocabulary acquisition. Hence, it is very important to have a clear goal in the study. There are tests devoted to measure vocabulary size (breadth), which focuses on the different frequency range of words. However, Nation (2016 cited by Read, 2019, p.546) distinguishes two types of vocabulary size tests: those that estimate the total vocabulary size and those –level tests– that measure how many words of a specific range are known. Tests measuring vocabulary breadth are the most common in SLA literature, although other options exist. There are also tests measuring vocabulary outcomes after a specific training, tests measuring vocabulary learning skills or those studying the cognitive process involved in vocabulary acquisition. Besides, there are tests designed to measure receptive or productive vocabulary whether spoken or written as well as those testing vocabulary depth (Read, 2019).

Regarding the format, the options are more limited. Tests can be group-administered or individual tests; they can be paper-based or computer-based as well as spoken or written. In relation to the inner structure of tests, there are also some possibilities like yes/no tests, multiple-choice tests or translation tests, among others (Read, 2019). Research analysing the benefits and drawbacks of the different types of test format has been developed and Nation (2001) includes a very interesting chapter on this on his book *Learning Vocabulary in Another Language*. Thus, the countless possibilities available can make it difficult to choose the most appropriate test for a context. According to Nation (2001), “a good vocabulary test has plenty of items” (p.345) and includes items which require learners to use the type of vocabulary that is going to be assessed. It should also be easy to make and interpret, with useful effects for the learning and teaching practice.

In the present study, receptive vocabulary is measured in year 3 of primary education (8 to 9-year-olds) both in students’ L1 (Spanish) and L2 (English). In order to compare vocabulary knowledge in both languages, the same test had to be passed in each language. In addition, selecting the most suitable test was difficult due to the low proficiency level of students in English and their short age, which usually makes it challenging to keep them engaged in the task.

Considering all these aspects, the test chosen was PPVT-III PEABODY- *Test de Vocabulario en Imágenes*, Spanish version by Dunn, Dunn & Arribas (2006). This test measuring receptive vocabulary was created to assess vocabulary in the L1 although the authors state that it can also be used to measure vocabulary in a foreign language. It is an individual,

multiple-choice oral test, which can be used for people from 2 years and six months of age up to 90 years. According to Olabarrieta-Landa et al. (2017), PPVT is a common instrument for measuring L1 vocabulary and one of its biggest advantage is the short administration time (although it needs to be administered individually) and the little oral communication required to complete the test since the examinee can point to the answer. The first edition of the test was produced in the United States in 1959 and the authors have carried out different revisions throughout the years to reduce gender, race, cultural or religious biases. The most recent revision in English (PPVT-IV) dates from 2007 and has no equivalent in Spanish. The version used in the present study is the adapted version of the PPVT-III used in Spain and Latin America, although the standardization of the test was only done in Spain (Olabarrieta-Landa et al., 2017).

Many vocabulary tests like the PPVT-III have been designed to be used with native speakers. However, an L1 and an L2 are not learned in the same way and this is why using a test designed for native speakers with L2 learners can pose some difficulties. Considering this, some researchers have analysed the reliability of Peabody Test for children learning an L2. In their study, they claimed that PPVT is also widely used in the field of SLA. They also point out that “one of the domains in which the PPVT is frequently used with L2 learners is the domain of early foreign-language education” (Goriot et al., 2018, p. 2). Finally, they also argue that, to investigate receptive vocabulary gains in learners, the use of this test is more and more popular in European countries where English has become the language of instruction. However, students going through this type of education do not learn the L2 in the same way as they learn their L1. Thus, Goriot et al. (2018) state that:

Different factors, such as the frequency with which certain words are used in the L2 or the linguistic overlap between the L1 and L2, may play a role in their L2 vocabulary development and influence their scores on the PPVT. (Goriot et al., 2018, p. 2)

In the study that Goriot et al. (2018) carried out, the aim was to investigate if Peabody was suitable to measure receptive vocabulary of Dutch students learning English as an L2 at different stages. They conducted three experiments to investigate the influence of L1 word frequency, cognates and learners' age when completing the test. They obtained the expected results: L1 word frequency as well as cognate status were determining factors specially with younger learners. Finally, they concluded that PPVT was not suitable to use with very limited proficiency learners but it was useful to compare L2 vocabulary knowledge in children with the same L1 (Goriot et al., 2018).

8. The study

8.1. Participants

In the present study, one group of year three of primary education is analysed (8-9 years old). The group is made up of 25 students but only 19 of them could complete the tests (due to the Covid-19 lockdown). These students attend “Las Vaguadas” bilingual school in Extremadura. The school is situated in an urban area in the outskirts of the capital city of the region, Badajoz. Students are considered medium-high class and they are involved in the CLIL project where they learn some subjects through English. Most of the members of the group have been studying in the school since pre-primary. This means that they have been involved in a CLIL project for six years although CLIL is properly implemented in primary. Thus, it can be said that most students have been in a bilingual project for two and a half years. However, there are some exceptions. One student joined the school in the first year of primary education who had been previously attending a bilingual nursery. Another student joined the school this academic year, coming from South America where his exposure to English was very limited and he had no previous experience with CLIL. In addition, there is a bilingual student in the group whose father is from Australia. Most students in the group also attend English lessons after school since they started primary education, but they have never travelled abroad to practice English (except the bilingual student).

The exposure time to English in the CLIL project is high: natural science (2 hours per week), social science (2 hours per week) and arts and crafts (1 hour per week). In addition, they have English lessons (3 hours per week), which makes a total of 8 hours of instruction in English per week (34 hours per month). Considering that most students have attended the school since the beginning of primary education (24 months), it can be estimated that they have received around 768 hours of instruction in English at school (Table 2).

Table 2

Participants' information.

MEMBERS	19
AGE	8-9
HOURS OF INSTRUCTIONS	768

Table 2: Participants' information

Source: own source

8.2. Instruments

In this study, two instruments were used with different purposes:

- A background information questionnaire. A questionnaire to collect the participant's background information was first designed. It is a simple questionnaire with three questions for students to answer in relation to their exposure time to English. They were asked when they started in the bilingual school, whether they attended English lessons after school or whether they had travelled abroad with the purpose of learning English. The questionnaire was completed individually before the vocabulary test.
- Peabody: *Test de Vocabulario en Imágenes*, PPVT-III. The Spanish version was used for measuring the receptive vocabulary in Spanish and the test was then adapted into English using Cambridge Dictionary to measure the receptive vocabulary of the same students in English. According to the authors, the test can be used both to measure L1 receptive vocabulary and to measure foreign language vocabulary. The test consists of 192 sheets with four drawings in black and white in each sheet. It is organized in 16 sets with 12 items in each set, which are organized by age groups, from the youngest to the oldest. The examiner should know the candidate's age in order to choose the starting point. Once the starting set is known, the examiner verbally says the word corresponding to the sheet and the candidate must decide the picture that best fits to the word. If there are 2 or more mistakes within a set, the examiner must go to the previous set until the candidate makes 0 or 1 mistake in a specific set. The test is finished when the candidate makes 8 or more mistakes within the same set.

8.3.Data collection

The tests were passed in five different days, from the 9th to the 13th of March 2020. The tests were completed individually in a quiet room next to the classroom. First, students completed the background questionnaire and then the English test. Once all students had finished the English version, the Spanish test was completed one student at a time. This was the procedure used because PPVT-III was designed to test students individually. In addition, students had two days between the English and Spanish test in order to avoid the testees, given their short age, to get tired or bored. Before the individual sessions, students were encouraged to complete the tests the best they could, and the tests were presented as a very attractive activity to do. For this reason, all students were highly motivated to complete the tests and they felt excited when it was their turn.

In the English version, all students started with the first set of words, which does not correspond to their chronological age, but which were more appropriate to their level of English. However, most of them covered several sets of words and, as a result, the English test took longer than expected to be completed. In the Spanish version, all students started in the set of words corresponding to their age. Therefore, they took the expected time, approximately 10 to 15 minutes per children.

Once all the tests were completed, each test was given the score, i.e., the number of the highest word reached minus number of errors. Then, the data was transcribed to two Excel documents. In the first Excel, student's scores and their grades in different subjects were collected. In the second Excel, the different words and the number of students who guessed each word were gathered. From this data, two different analyses were carried out. First, a quantitative analysis comparing scores in both tests and their relation to the students' marks. Then, a qualitative analysis taking into account which words most students know or do not know and the different features of each group of words.

8.4.Results and discussion

In this section, the results of the study are presented in relation to the different research questions stated in this MA dissertation. The data of the first and second research questions are analysed using quantitative measurements in SPSS, more specifically, a Spearman correlation. However, the third research question is analysed qualitatively. As a means to understand better the results, first the sample description is presented. There are 19 sample subjects and the information about them includes: Peabody test score in in English and Spanish, marks or grade in the following subjects: English, natural science, social science and Spanish. The descriptive statistical information can be seen in Table 3.

Table 3

Descriptive Statistics.

	N	Min.	Max.	Average	Standard dev.	Variance
English test (ET)	19	15	94	64.42	22.282	496.480
Spanish test (ST)	19	97	133	109.26	8.130	66.094
English marks (EM)	19	5.00	10.00	8.7895	1.54844	2.398
Natural science marks (NSM)	19	4.00	10.00	8.5263	1.64548	2.708
Social science marks (SSM)	19	4.00	10.00	8.1579	1.74047	3.029
Spanish marks (SM)	19	7.00	10.00	8.2632	.99119	.982
N	19					

Table 3: Descriptive Statistics

Source: own source

To answer research questions number one and two, a correlation test was performed, although a test of normality was required to decide on the specific statistical test. As most variables show a non-normal distribution (see in the Appendices), a nonparametric correlation test is chosen, using Spearman's coefficient. Results of the correlation study are presented in Table 4.

Table 4
Correlations.

			ET	ST	EM	NSM	SSM	SM
Spearm an's rho	English Test (ET)	Correlation Coefficient	1.000	.326	.667(**)	.296	.707(**)	.550(*)
		bilateral sigma	.	.173	.002	.219	.001	.015
		N	19	19	19	19	19	19
	Spanish Test (ST)	Correlation Coefficient	.326	1.000	.038	.162	.207	.050
		bilateral sigma	.173	.	.878	.508	.396	.839
		N	19	19	19	19	19	19
	English Mark (EM)	Correlation Coefficient	.667(**)	.038	1.000	.689(**)	.831(**)	.784(**)
		bilateral sigma	.002	.878	.	.001	.000	.000
		N	19	19	19	19	19	19
	Natural Science Mark (NSM)	Correlation Coefficient	.296	.162	.689(**)	1.000	.568(*)	.717(**)
		bilateral sigma	.219	.508	.001	.	.011	.001
		N	19	19	19	19	19	19
	Social Science Mark (SSM)	Correlation Coefficient	.707(**)	.207	.831(**)	.568(*)	1.000	.703(**)
		bilateral sigma	.001	.396	.000	.011	.	.001
		N	19	19	19	19	19	19
	Spanish Mark (SM)	Correlation Coefficient	.550(*)	.050	.784(**)	.717(**)	.703(**)	1.000
		bilateral sigma	.015	.839	.000	.001	.001	.
		N	19	19	19	19	19	19

Table 4: Correlations

Source: own source

8.4.1. RQ 1: Do children in 3rd grade of primary education with a high score of receptive vocabulary in their L1 also have a high score of receptive vocabulary in their L2?

In this research question, the main aim is to analyse the relationship between scores in the English test and the Spanish test. A correlation coefficient was computed to assess the relationship between scores in the English and the Spanish Peabody test. As can be appreciated from Table 3, scores in English are lower than scores in the Spanish test. In addition, the standard deviation is much higher in the English tests, meaning that students obtained more varied scores in the English test than in the Spanish test. The distribution of the students in relation to their score obtained in English and in Spanish can be clearly appreciated in Figure 1.

Figure 1

Distribution of students according to their scores in the tests.

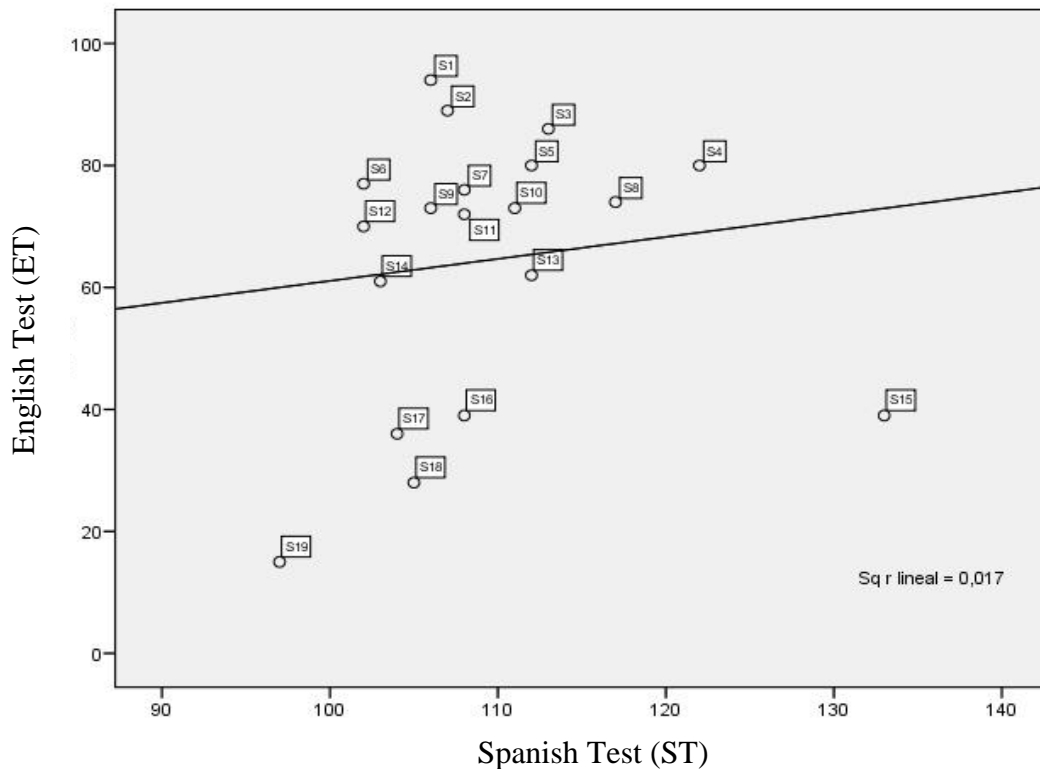


Figure 1: Distribution of students according to their scores in the tests

Source: own source

Figure 1 shows that in this study getting a higher score in the Peabody Spanish test does not correlate with getting a higher score in the Peabody English test. We can see a moderately upward slope, but the results are not near the regression line. This is also noticeable in the statistics results (see Table 4), where it can be observed that scores in the Peabody Spanish test do not show a statistically significant correlation with scores in the Peabody English test ($r = .326$, $n = 19$, $p = .173$). It can also be observed that scores in the Peabody Spanish test are more homogeneous than scores in the Peabody English test (Table 3). The maximum score in Spanish is 133 and the minimum is 97 (standard deviation 22.3), whereas in the Peabody English test the maximum is 96 and the minimum is 15 (standard deviation 8.1). To a lesser extent, this phenomenon can be also appreciated in students' marks where the heterogeneity among marks in subjects taught in English is higher than in Spanish marks. In English, the maximum mark is a 10 and the minimum is a 5 (standard deviation 1.5), in natural and social science the minimum is a 4 and the maximum is a 10 (standard deviation 1.6 and 1.7 respectively). However, the maximum mark in Spanish is a 10 and the minimum is a 7 (standard deviation .99).

As discussed in previous sections, experts on the field agree (e.g., Webb & Nation, 2017), that L1 vocabulary grows faster than L2 vocabulary. Therefore, the fact that students got higher scores in the Spanish test was expected, as this is their L1.

Considering the heterogeneity of scores in the Peabody English test in comparison with the scores in the Peabody Spanish test, one possible explanation could be the presence of a new student (S19) in the group coming from South America. He had recently joined a CLIL school for the first time and English as a subject had not been very common in his school experience. In Figure 1, it can be seen that S19 got the lower score in the English test, which is very distant to the highest score obtained by S1. In addition, S1 is a bilingual student (Spanish mother and Australian father) who learned both languages simultaneously since birth and whose contact with both languages is usual in his everyday life. However, it can be pointed out that another student, S2, got similar scores to S1 in the Peabody English test. One possible reason for this is the fact that S2 had attended a bilingual nursery, where only English is spoken by the teachers who are native speakers of English, since she was a baby until the moment she started primary education.

In this research question, it was expected, as Meara (1996) had already found, that students with higher scores in the Spanish test would also get higher scores in the English test. However, it can be clearly seen from Table 4 that no correlations were found between scores in the

Spanish and the English test in this study. As proposed by Leśniewska & Pichette (2016), the explanation might be related to the importance of individual differences and environmental influences in the students' L2 backgrounds.

8.4.2. RQ 2: Are the scores in L1 vocabulary related to students' marks in the content subjects taught in Spanish? Are scores in L2 vocabulary related to students' marks in the content subjects taught in English?

This research question aims to analyse, on the one hand, the relation between students' scores in the Peabody Spanish test and their marks in content subjects taught in Spanish and, on the other, students' scores in the Peabody English test and their marks in the content subjects taught in English. To study this, a correlation was carried out, as can be seen from Table 4.

The results obtained show that the Peabody scores in English highly correlate with the marks or grades in social science ($r = .707^{**}$, $n=19$, $p = .001$) and in second place with the marks in the EFL subject ($r = .667^{**}$, $n=19$, $p = .002$). To a lesser extent, the Peabody English test scores correlate with the mark in Spanish as a school subject ($r = .550^*$, $n=19$, $p = .015$) but it is not significantly correlated with the mark in natural science. As far as the Peabody test scores in Spanish are concerned, no correlation with any of the subject marks can be found. However, the Spanish mark presents high correlation with the marks in the other three subjects: English ($r = .784^{**}$, $n=19$, $p = .000$), natural science ($r = .717^{**}$, $n=19$, $p = .001$) and social science ($r = .703^{**}$, $n=19$, $p = .001$). Regarding the English marks, apart from the correlation with the Peabody English test, it is also correlated with the rest of the marks. The highest correlation is with the social science mark ($r = .831^{**}$, $n=19$, $p = .000$), followed by the Spanish mark ($r = .784^{**}$, $n=19$, $p = .000$) and finally with the natural science mark ($r = .689^{**}$, $n=19$, $p = .001$). Finally, the natural and social science mark are also correlated between them ($r = .568^*$, $n=19$, $p = .011$).

In short, English scores in the Peabody test are correlated with marks in social science, English and Spanish (in this order) and not with marks in natural science. Spanish scores in the Peabody test are not correlated with any of the items. However, although it was not part of the research question, marks in Spanish are correlated with marks in English, natural science and social science and marks in English are correlated with marks in social science, Spanish and natural science (in this order).

Table 5

Correlation between students' vocabulary and their marks (summary).

	English Mark (EM)	Natural Science Mark (NSM)	Social Science Mark (SSM)	Spanish Mark (SM)
English Test (ET)	.667**	---	.707**	.550*
Spanish Test (ST)	---	---	---	---

Table 5: Correlation between students' vocabulary and their marks (summary)

Source: own source

It was expected that results in both tests were correlated with students' marks although results show that only the Peabody English test is correlated with some marks. Conversely, the Peabody Spanish test is not correlated with any of the variables, which was not anticipated.

To start with, the Peabody English test correlates with the marks in social science as well as with the English marks. However, it is noticeable that there is no correlation between the Peabody English test and the natural science mark.

The highest correlation is between the Peabody English test and the social science mark (this correlation is also noticeable between the English mark and the social science mark). This might be related to the fact that social science is a dense subject where information is packed and abstract, which makes it more difficult for students. However, natural science seems to be an easier subject for students because information is more concrete and visual, thus more accessible for students. The same was observed in a study conducted by Píriz Rico (2015) in 5th grade in a CLIL school in Extremadura where students expressed the same feeling. The correlation between the Peabody English test with the social science mark might mean that students with better level of vocabulary are able to achieve better results in social science.

The correlation between students' vocabulary and their marks was pointed out by Webb & Nation (2017). The Peabody English test is also correlated with the Spanish mark, so it can be assumed that getting higher scores in the vocabulary test in English, i.e. having a higher level of vocabulary in English, is not affecting negatively students' marks in their L1.

As it was pointed out before, the Peabody Spanish test is not correlated with any other variable. Although correlations between the Spanish marks (as a school subject) and the other marks are remarkable, there is no evidence that vocabulary is responsible of this correlation.

8.4.3. RQ 3: Is it possible to identify a core L2 vocabulary for 3rd graders, i.e. a group of English words known by most students at this educational stage versus a peripheral vocabulary, i.e. a group of words only known by some students? What are the different features that each of these two groups of words share?

The aim of this research questions is to analyse whether there is a core vocabulary known by most students (words that most students answered correctly) and a peripheral vocabulary, i.e. words that most students did not answer correctly. The final aim is to analyse different features that could be analysed in each group of words, e.g. cognates or exposure to input (words that had been studied).

8.4.3.1. Identifying the groups of words

To analyse words, what we have termed core and peripheral vocabulary are first identified. In order to do this, terciles were calculated. The total amount of words that students answered in English was 120 words. They were divided into three different groups, each group containing a third of the words. This resulted in three groups of around 40 words. The group of words with low percentage of correct answers (from 0% to 42.86%) is considered the peripheral vocabulary because only a minority of the students knew them. Then, there is a middle group of words with medium percentage of correct answers (from 44.44% to 84.21%). Finally, the core vocabulary is the group of words with high percentage of correct answers (84.62% to 100%). The distribution of words can be seen in Table 5 in the Appendices.

Taking into account the distinction between words with low, medium and high percentage of correct answers, it can be appreciated that these words are distributed throughout the test, which is organized in sets of words depending on age (from young to old). However, it can be noted that in the first sets there are fewer difficult words (words from the peripheral group) which increase with the different sets (see Table 6 in the Appendices). This is a sign that the words are suitably distributed, reinforcing the validity of the test.

8.4.3.2. Core vocabulary

When analysing the words, it can be appreciated that within the core vocabulary, most of the words known by all students (100%) which are: *lamp, helicopter, accident, kangaroo, fruit, cactus, astronaut, and archer* are all cognates which are pronounced in a similar way in both languages (orthography is not important as students are only exposed to the oral input). In this case, no information about students' exposure to these words in textbooks or oral input was available. The fact that cognates could be a factor that affected students' vocabulary knowledge was expected, as it is explicitly mentioned in the study carried out by Goriot et al. (2018) where the Peabody test is used and the influence of cognates is studied. This is not surprising as Spanish-English cognates are a common reality (there are around 10,000-15,000 Spanish-English cognates) and this is beneficial especially for CLIL students as a high percentage of English academic words are Spanish-English cognates which are very frequent in Spanish (Lubliner & Hiebert, 2011).

There are other words that got 100% of correct answers. These words are *welcome, cliff, island, brain, and writing*. In this case, there are no Spanish-English cognates among these words. However, students' exposure to these words in school is high. The words *welcome* and *writing* are used every day in class routines while the words *cliff, island* and *brain* are technical words studied in social and natural science. The students admitted that the word *brain* was especially familiar for them, as it was part of their last natural science exam. It is worth mentioning that the word *welcome* got higher percentage of correct answers in English than in Spanish (*acoger*).

Inside the core vocabulary, there are other words that most students answered correctly which can be highlighted. First, we can also find Spanish-English cognates in this group of words: *globe, calculator, rectangle, trunk, classify, calculate, predator, infinity, instruct, oval, statue, reptile, palm, and composer*. Some of these words can be also considered academic words which had been studied by students in the CLIL subjects. For example, most students noted that words like *globe, calculator, predator, and reptile* were very easy because they had studied them in natural science. Then, within core vocabulary other words can be found like *fox, porcupine, pollute, oil, mammal, and vegetable*. These words are not cognates, but they can be considered as academic words for CLIL students who had been exposed to this terminology in natural science and they are familiar with them. From these words that most students answered correctly, it can be pointed out that the words *classify (clasificar), instruct (instruir), oil (lubricar), and mammal (mamífero)* got higher percentage of correct answers in

English than in Spanish. This could be related to a higher exposure to specific vocabulary related to content subjects in English.

Finally, words like *drink*, *money*, *climb*, *cow*, *swim*, and *throw* can be considered high-frequency words in English. To sum up, in core vocabulary three main features can be found: cognates, exposure to input and high-frequency words.

8.4.3.3. Peripheral vocabulary

Considering peripheral vocabulary, i.e., words that most students do not know, different words can be found. First, words that no student answered correctly are analysed. These words are *frame*, *bouquet*, and *glider*. These words are not Spanish-English cognates and probably their exposure is not very high. It is remarkable that the Spanish version of the word *glider* (*planeador*) got very low percentage of correct answers too (26.32%). The word *bouquet* (*ramo*) got higher percentages in Spanish (57.89%) while the word *frame* was in the first sets of the test and students were not required to answer in Spanish.

There are other words that very few students answered correctly. Some of these words are *jaw*, *pod*, *moan*, *head* (verb), *shovel*, *celery*, *hairy*, *harvest*, *empty* or *crash*. It is remarkable that none of these words is a Spanish-English cognate, which again supports Goriot et al. (2018) findings. However, there are interesting aspects to comment. The word *shovel*, which was one of the first words to appear in the test, can be considered a very frequent word for native children, as it is very popular in young children games. Nevertheless, it is very infrequent for L2 learners who are not exposed to that type of contexts in the L2. In this case, only 3 children answered correctly including the bilingual student and the student who attended a bilingual nursery for 6 years.

Moreover, words like *jaw* (*mandíbula*) and *head* (*dirigir*) got very high percentages of correct answers in Spanish. This may suggest that although the vocabulary is clear in the L1, those words are not part of most students' L2 vocabulary. The word *head* is recognized by most students when it is used as a noun but not as a verb. In a similar way, *moan* and *harvest* also got higher percentage in Spanish. On the other hand, the word *pod*, did not get very high percentages in Spanish. In addition, words like *celery*, *hairy* and *empty* cannot be compared to percentages in Spanish as they appear in the first sets where students were not required to answer.

Furthermore, there are words within the peripheral vocabulary that got higher percentages of correct answers in English than in Spanish. This is the case of words like *confidential*, *valve*, *wedge*, and *crash*. A possible explanation for this can be related to the reduced number of students that answered to these words in English. In fact, the number of students that answered correctly is very similar in both languages. The difference is the sample which makes that results, in terms of percentages, are different. The reason for the reduced number of students that answered in English is related to the test's rules. When students make 8 or more mistakes within a set, they cannot continue answering the following sets of words. The percentages of correct answers in English and Spanish of the different words can be seen in Table 7 in the Appendices.

From this research question, it can be concluded that cognates are an important feature when comparing L1 and L2 vocabulary as Goriot et al. (2018) state in their study. In addition, exposure to input should also be considered as Castellano-Risco (2018) concluded in her research.

9. Conclusions

Once the data from the three research questions have been analysed and discussed, some conclusions can be drawn:

1. Importance of studying different variables affecting vocabulary knowledge: In the present study, no correlation was found between students' L1 and L2 vocabulary. This finding does not match what Meara (1996) and Miralpeix (2019) reveal in their work: despite individual differences, good vocabulary learners in an L1 are usually good vocabulary learners in an L2. However, similar results were found in a study conducted by Leśniewska & Pichette (2016) where having a larger lexicon in the L1 was not correlated with being good at learning L2 vocabulary. Thus, it can be concluded that when researching about vocabulary knowledge it is important to take into account learners' individual differences and environmental influences that may affect students' vocabulary knowledge in an L1 and an L2.
2. Benefits of L2 vocabulary knowledge is students' marks: Correlation was found between students' L2 vocabulary and students' marks, as it has been stated by Laufer & Goldstein (2004). This is especially important in CLIL where learners study different content subjects through a second language. The fact that social science was found to be the subject where correlation with L2 vocabulary was clearer, may suggest the importance of vocabulary knowledge to understand packed and dense information. Explanations in social science tend to be more linguistic and abstract and there are not tangible experiments to demonstrate phenomena. In addition, contrary to a general belief in the educational community, in this study, larger lexical knowledge in an L2 is also positively correlated with students' marks in their L1. On the other hand, no correlation was found between students' L1 vocabulary and students' marks. This absence of correlation might be related to the fact that only one subject taught in students' L1 was taken into account when computing the correlation.
3. The importance of cognates: When analysing the core and peripheral vocabulary, it could be noticed that Spanish-English cognates are easier for students. This is not surprising as several researchers have claimed the same, e.g. Goriot et al., (2018); Lubliner & Hiebert, (2011); Webb & Nation, (2017). Considering exposure to input, in the present study is not analysed deeply. Specific information about some words reveal the positive influence of exposure to input when learning vocabulary.

However, as it was claimed by Castellano-Risco et al. (2020), it is not the amount of input but the educational context that CLIL learners are exposed to that can offer advantages for vocabulary knowledge.

Everything considered, it can be concluded that vocabulary is a key aspect in second language acquisition which can have great benefits for learners. This is extremely important in CLIL settings where students learn content through a second language. The significance of this is related to the fact that students need to have high lexical competence in the L2 in order to access content. Here we need to point out the importance of CLIL teachers being aware of their students' vocabulary knowledge and how to help them in this difficult task which has been traditionally underestimated.

Some important implications for teaching can be highlighted, which can be specially interesting for CLIL primary school teachers:

1. Importance of L2 vocabulary in CLIL: As Meara (1996) states, "most language teachers are remarkably ill informed about the role that lexis plays in language". (p. 36). CLIL teachers should be aware that teaching grammar and vocabulary separately does not make much sense. As Lewis (1993) claims, language consists of grammaticalized lexis, meaning that learning vocabulary within a context and structures is crucial. This is especially important in CLIL where learners are learning language and content at the same time (close to a natural setting). Thus, information about how to use words appropriately within a specific context is needed. In addition, this study shows that L2 vocabulary influences positively students' L2 and L1 marks. This is a tool that CLIL teachers can use to help their students improve their learning process.
2. Importance of research to improve teaching practices: There are no better tools to know how to improve the teaching-learning process than research. It is only with empirical evidence that teachers will know the influence of vocabulary in their students' learning and how to exploit the advantages.
3. Importance of vocabulary in social science versus natural science: In light of the results obtained, CLIL teachers should be aware of the differences between natural science and social science. Even though they might seem very similar, results show that they are not only different content-wise but also in relation to language. The importance of vocabulary in social science should be taken into account to help students understand packed and abstract information. On the other hand, natural science can be more visual and concrete which makes it easier for students. For this

reason, teachers should not only teach technical vocabulary which might be very obvious in natural science, but also academic language which will facilitate the understanding of dense information.

Finally, the present study has some limitations that are worth mentioning:

1. **Size of sample:** The sample was thought to be a group of 25 students but due to the global pandemic caused by the covid-19 only 19 students could answer the test, as classes got interrupted. In addition, it would be a good option to conduct this study to a wider sample, e.g. all groups of 3rd grade in the school (75 students) and to do it in other schools with different characteristics.
2. **Studying vocabulary development:** It would be interesting to study vocabulary growth in 3rd grade of primary education with a pre-test and a post-test in order to gather information about students' improvement in one academic year.
3. **Studying other variables:** In this study receptive vocabulary in an L1 and an L2 is compared. However, it would be a good idea to study other variables like gender or learners' strategies when learning vocabulary. Furthermore, comparing receptive vocabulary between CLIL and EFL learners in a first and second language would be very enriching to see the learning outcomes of both approaches in terms of vocabulary acquisition.
4. **Studying students' exposure to input in CLIL lessons:** Taking a closer look at the input that students are exposed to would be also interesting data to consider. Not only in terms of amount of input but also quality of input from the teacher, or textbooks.
5. **Detailed qualitative analysis of words:** A more detailed of the core and peripheral vocabulary would have been of great interest for further research on vocabulary acquisition. There are some aspects like word frequency in English and Spanish, part of speech and CEFR level that would contribute to develop a very enriching study of the different features that can affect vocabulary knowledge.
6. **Comparison with other studies:** It is always advisable to compare results with studies of similar characteristics. In this case, studies about vocabulary acquisition have been contrasted but there are few of them with matching characteristics. It would be interesting to compare this study with others using PPVT with learners sharing the same mother tongue, i.e., Spanish and the same second language, i.e., English.

10. References

- Agustín-Llach, M. P., & Canga Alonso, A. (2016). Vocabulary growth in young CLIL and traditional EFL learners: evidence from research and implications for education: Vocabulary growth in young CLIL and traditional EFL learners. *International Journal of Applied Linguistics*, 26(2), 211–227. <https://doi.org/10.1111/ijal.12090>
- Alejo, R., & Piquer-Píriz, A. M. (2010). CLIL teacher training in Extremadura: A need analysis perspective. In D. Lasagabaster & Y. Ruiz de Zarobe, *CLIC in Spain: Implementation, Results and Teacher Training*. (pp. 219–242). Cambridge Scholars Publishing.
- Bruton, A. (2011). Is CLIL so beneficial, or just selective? Re-evaluating some of the research. *System*, 39(4), 523–532. <https://doi.org/10.1016/j.system.2011.08.002>
- Canga Alonso, A. (2015). Receptive Vocabulary of CLIL and Non-CLIL Primary and Secondary School Learners. *Complutense Journal of English Studies*, 23(0), 59–77. https://doi.org/10.5209/rev_CJES.2015.v23.51301
- Castellano-Risco, I. (2018). Receptive Vocabulary and Learning Strategies in Secondary School CLIL and non-CLIL Learners. *Onomázein Revista de Lingüística Filología Y Traducción*, 40, 28–48. <https://doi.org/10.7764/onomazein.40.02>
- Castellano-Risco, I., Alejo-González, R., & Piquer-Píriz, A. (2020). The development of receptive vocabulary in CLIL vs EFL: Is the learning context the main variable? *System*, 102263. <https://doi.org/10.1016/j.system.2020.102263>
- Cenoz, J., Genesee, F., & Gorter, D. (2014). Critical Analysis of CLIL: Taking Stock and Looking Forward. *Applied Linguistics*, 35(3), 243–262. <https://doi.org/10.1093/applin/amt011>
- Cook, V. (2008). *Second language learning and language teaching* (4. ed). Routledge.
- Coyle, D., Hood, P., & Marsh, D. (2010). A window on CLIL. In D. Coyle, P. Hood, & D. Marsh, *CLIL: Content and Language Integrated Learning*. (pp. 1–13). Cambridge University Press.
- Dunn, L. M., Dunn, L. M., & Arribas, D. (2006). *PPVT-III Peabody: test de vocabulario en imágenes*. Tea.
- Ellis, N., & Beaton, A. (1993). Factors Affecting the Learning of Foreign Language Vocabulary: Imagery Keyword Mediators and Phonological Short-Term Memory. *The Quarterly Journal of Experimental Psychology Section A*, 46(3), 533–558. <https://doi.org/10.1080/14640749308401062>

- Gablasova, D. (2014). Learning and Retaining Specialized Vocabulary From Textbook Reading: Comparison of Learning Outcomes Through L1 and L2. *The Modern Language Journal*, 98(4), 976–991. <https://doi.org/10.1111/modl.12150>
- Goriot, C., van Hout, R., Broersma, M., Lobo, V., McQueen, J. M., & Unsworth, S. (2018). Using the Peabody picture vocabulary test in L2 children and adolescents: effects of L1. *International Journal of Bilingual Education and Bilingualism*, 1–23. <https://doi.org/10.1080/13670050.2018.1494131>
- Jiménez Catalán, R. M., & Ruiz de Zarobe, Y. (2009). The receptive vocabulary EFL learners in two instructional contexts: CLIL versus non-CLIL instruction. In R. M. Jiménez Catalán & Y. Ruiz de Zarobe (Eds.), *Content and language integrated learning: evidence from research in Europe* (pp. 81–92). Multilingual Matters.
- Language. (n.d.). In *Macmillan Dictionary*. <https://www.macmillandictionary.com/dictionary/british/language>
- Laufer, B., & Goldstein, Z. (2004). Testing Vocabulary Knowledge: Size, Strength, and Computer Adaptiveness: Language Learning. *Language Learning*, 54(3), 399–436. <https://doi.org/10.1111/j.0023-8333.2004.00260.x>
- Leśniewska, J., & Pichette, F. (2016). Songs vs. stories: impact of input sources on ESL vocabulary acquisition by preliterate children. *International Journal of Bilingual Education and Bilingualism*, 19(1), 18–34. <https://doi.org/10.1080/13670050.2014.960360>
- Lewis, M. (1993). *The lexical approach: the state of ELT and a way forward* (6. [print.]). Thomson Heinle.
- Lightbown, P. M., & Spada, N. (2013). *How languages are learned*. (4th Edition). Oxford University Press.
- Li, M., & Kirby, J. R. (2014). The Effects of Vocabulary Breadth and Depth on English Reading. *Applied Linguistics*, 36/5, 611–634. <https://doi.org/10.1093/applin/amu007>
- Lindgren, E., & Muñoz, C. (2013). The influence of exposure, parents, and linguistic distance on young European learners' foreign language comprehension. *International Journal of Multilingualism*, 10(1), 105–129. <https://doi.org/10.1080/14790718.2012.679275>
- Lublinter, S., & Hiebert, E. H. (2011). An Analysis of English–Spanish Cognates as a Source of General Academic Language. *Bilingual Research Journal*, 34(1), 76–93. <https://doi.org/10.1080/15235882.2011.568589>

- Meara, P. (1996). The dimension of lexical competence. In G. Brown, K. Malmkjaer, & J. Williams (Eds.), *Performance and competence in second language acquisition* (pp. 33–51). Cambridge University Press.
- Mehisto, P., Marsh, D., & Frigols, M. J. (2008). Approaching CLIL. In P. Mehisto, D. Marsh, & M. J. Frigols, *Uncovering CLIL: content and language integrated learning in bilingual and multilingual education*. (pp. 9–23). Macmillan.
- Merikivi, R., & Pietilä, P. (2014). Vocabulary in CLIL and in Mainstream Education. *Journal of Language Teaching and Research*, 5(3), 487–497. <https://doi.org/10.4304/jltr.5.3.487-497>
- Miralpeix, I. (2019). L1 and L2 Vocabulary Size and Growth. In S. A. Webb, *The Routledge handbook of vocabulary studies* (pp. 189–206). New York, NY : Routledge.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.
- Olabarrieta-Landa, L., Rivera, D., Ibáñez-Alfonso, J. A., Albaladejo-Blázquez, N., Martín-Lobo, P., Delgado-Mejía, I. D., Lara, L., Rabago Barajas, B. V., Rodriguez Salgado, A. M., Paredes Quispe, L. A., Romero-García, I., Velázquez-Cardoso, J., García de la Cadena, C., Fernandez-Agis, I., Padilla-López, A., Hernández Agurcia, G. P., Marín-Morales, A., Corral San José, A., & Arango-Lasprilla, J. C. (2017). Peabody Picture Vocabulary Test-III: Normative data for Spanish-speaking pediatric population. *NeuroRehabilitation*, 41(3), 687–694. <https://doi.org/10.3233/NRE-172239>
- Pérez Cañado, M. L. (2012). CLIL research in Europe: Past, present and future. *International Journal of Bilingual Education and Bilingualism*, 15(3), 315–341.
- Pérez Cañado, M. L. (2018). CLIL and Educational Level: A Longitudinal Study on the Impact of CLIL on Language Outcomes. *Porta Linguarum*, 29, 51–70.
- Pignot-Shahov, V. (2012). Measuring L2 receptive and productive vocabulary knowledge. *Language Studies Working Papers*, 4(1), 37–45.
- Píriz Rico, B. (2015). *Unfolding individual differences in the CLIL primary classroom: comparing 'social science' and 'natural science'. A study of motivation, anxiety and willingness to communicate in 5th grade in Extremadura*. [Universidad de Extremadura]. <http://dehesa.unex.es/handle/10662/3469>
- Read, J. (2019). Key Issues in Measuring Vocabulary Knowledge. In S. Webb (Ed.), *The Routledge Handbook of Vocabulary Studies*. (pp. 545–560). New York, NY : Routledge.

- Schleppegrell, M. J. (2006). The challenges of academic language in school subjects. In I. Lindbers & Sandwall (Eds.), *Språket och kunskapen: att lära på sitt andraspråk i skola och högskola*. (pp. 47–69). Göteborg universitet institutet för svenska som andrasprak.
- Schmitt, N., & Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching*, 47(4), 484–503.
<https://doi.org/10.1017/S0261444812000018>
- Second Language Acquisition. (n.d.). In *Macmillan Dictionary*.
<https://www.macmillandictionary.com/dictionary/british/second-language-acquisition>
- Webb, S. (2008). Receptive and productive vocabulary sizes of L2 learning. *Studies in Second Language Acquisition*, 30(1), 79–95.
- Webb, S. A., & Nation, I. S. P. (2017). *How vocabulary is learned*. Oxford University Press.
- Zhou, S. (2010). Comparing Receptive and Productive Academic Vocabulary Knowledge of Chinese EFL Learners. *Asian Social Science*, 6(10).
<https://doi.org/10.5539/ass.v6n10p14>

11. Appendices

Figure 2

Test of normality English Test

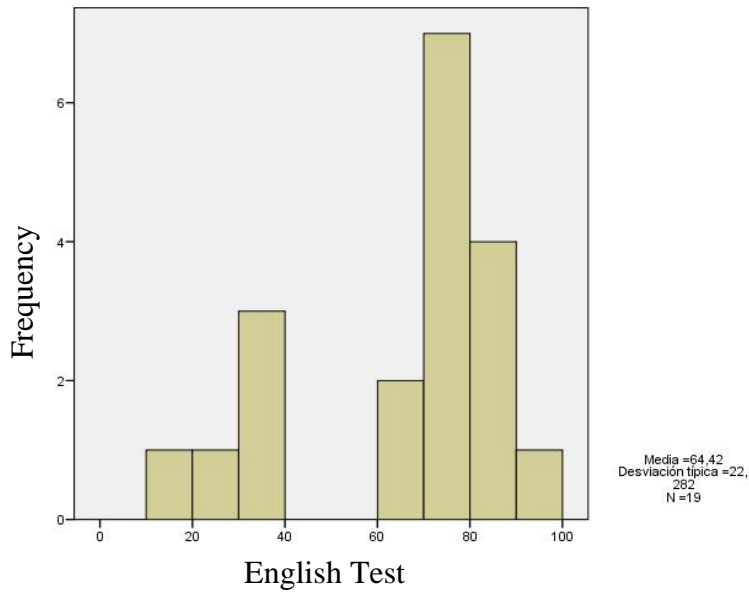


Figure 2: Test of normality English Test

Source: own source

Figure 3

Test of normality Spanish Test

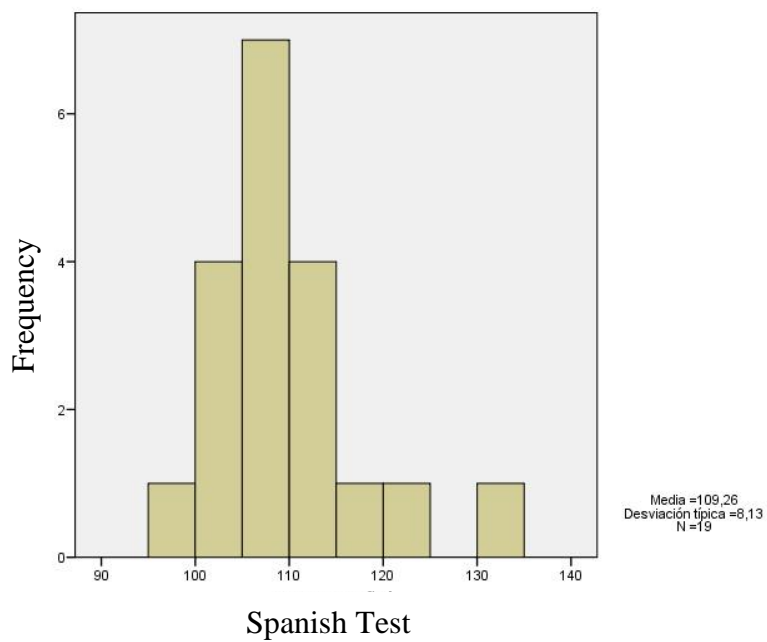


Figure 3: Test of normality Spanish Test

Source: own source

Figure 4

Test of normality English Mark

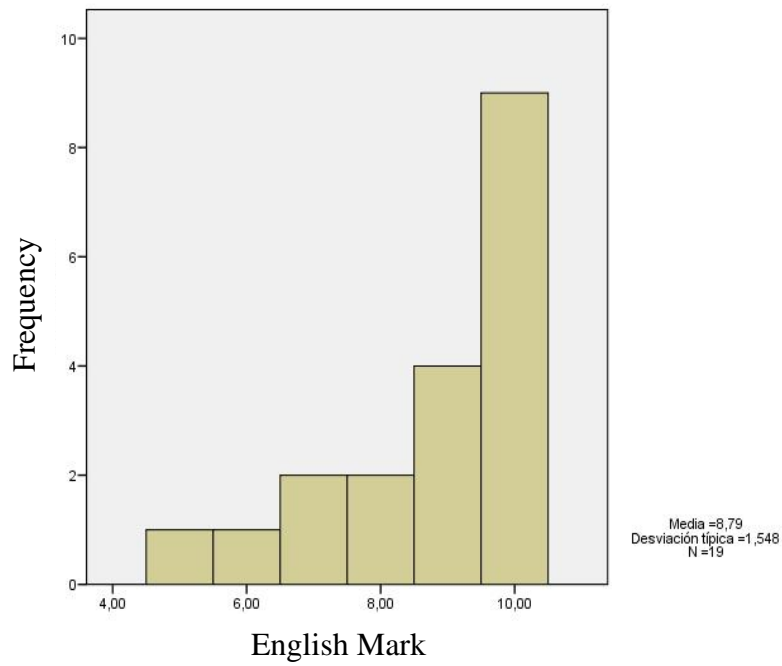


Figure 4: Test of normality English Mark

Source: own source

Figure 5

Test of normality Natural Science Mark

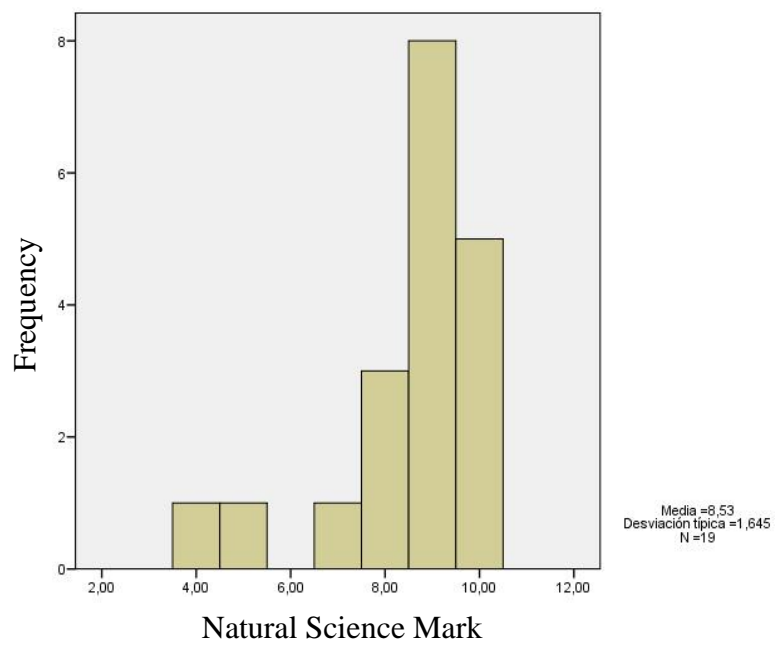


Figure 5: Test of normality Natural Science Mark

Source: own source

Figure 6

Test of normality Social Science Mark

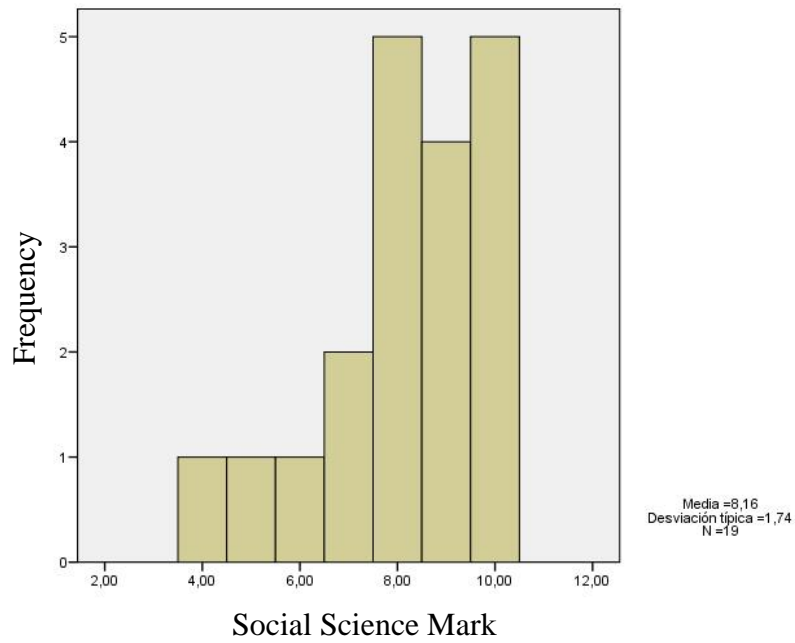


Figure 6: Test of normality Social Science Mark

Source: own source

Figure 7

Test of normality Spanish Mark

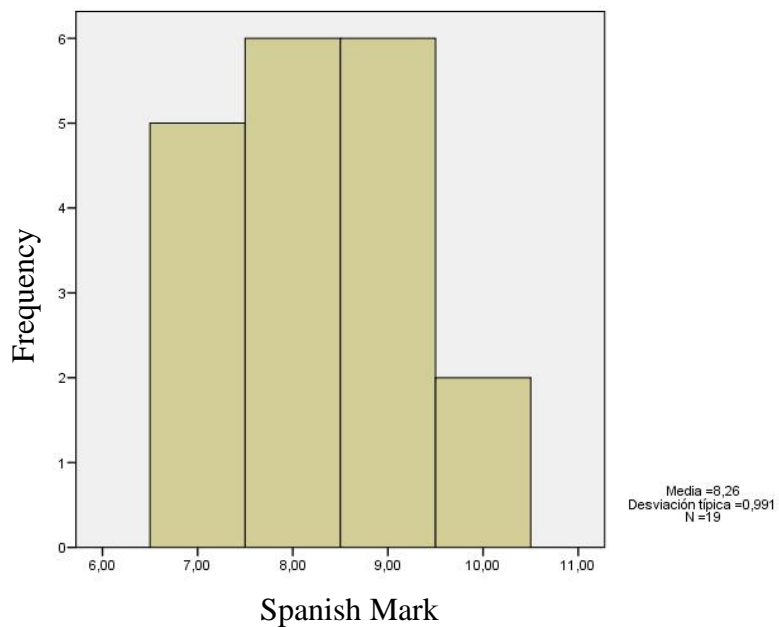


Figure 7: Test of normality Spanish Mark

Source: own source

Table 6

Distribution of words in English.

SET	Words	N English	English correct	% of correct answers	Frequency bands
SET 01	Broom	19	14	73,68%	Medium
SET 01	Plane	19	16	84,21%	Medium
SET 01	Drink	19	18	94,74%	High
SET 01	Shovel	19	3	15,79%	Low
SET 01	Swing	19	14	73,68%	Medium
SET 01	Lamp	19	19	100,00%	High
SET 01	Money	19	18	94,74%	High
SET 01	Helicopter	19	19	100,00%	High
SET 01	Fence	19	12	63,16%	Medium
SET 01	Key	19	9	47,37%	Medium
SET 01	Drum	19	9	47,37%	Medium
SET 01	Climb	19	18	94,74%	High
SET 02	Cow	19	18	94,74%	High
SET 02	Swim	19	18	94,74%	High
SET 02	Empty	19	4	21,05%	Low
SET 02	Excavate	19	16	84,21%	Medium
SET 02	Farmer	19	13	68,42%	Medium
SET 02	Accident	19	19	100,00%	High
SET 02	Nest	19	11	57,89%	Medium
SET 02	Throw	19	17	89,47%	High
SET 02	Envelope	19	9	47,37%	Medium
SET 02	Castle	19	16	84,21%	Medium
SET 02	Measure	19	10	52,63%	Medium
SET 02	Kangaroo	19	19	100,00%	High
SET 03	Fruit	19	19	100,00%	High
SET 03	Chain	19	12	63,16%	Medium
SET 03	Cactus	19	19	100,00%	High
SET 03	Porcupine	19	18	94,74%	High
SET 03	Yawn	19	9	47,37%	Medium
SET 03	Goat	19	11	57,89%	Medium
SET 03	Fancy	19	7	36,84%	Low
SET 03	Fox	19	18	94,74%	High
SET 03	Claw	19	7	36,84%	Low
SET 03	Argue	19	10	52,63%	Medium
SET 03	Astronaut	19	19	100,00%	High
SET 03	Saw	19	13	68,42%	Medium
SET 04	Trunk	18	17	94,44%	High
SET 04	Huge	18	6	33,33%	Low

SET 04	Parachute	18	9	50,00%	Medium
SET 04	Deliver	18	7	38,89%	Low
SET 04	Globe	18	17	94,44%	High
SET 04	Calculator	18	17	94,44%	High
SET 04	Drip	18	8	44,44%	Medium
SET 04	Hive	18	9	50,00%	Medium
SET 04	Sand	18	8	44,44%	Medium
SET 04	Statue	18	16	88,89%	High
SET 04	Terrified	18	11	61,11%	Medium
SET 04	Rectangle	18	17	94,44%	High
SET 05	Frame	18	0	0,00%	Low
SET 05	Baggage	18	14	77,78%	Medium
SET 05	Writing	18	18	100,00%	High
SET 05	Encourage	18	4	22,22%	Low
SET 05	Vehicle	18	4	22,22%	Low
SET 05	Polish	18	6	33,33%	Low
SET 05	Celery	18	3	16,67%	Low
SET 05	Oval	18	16	88,89%	High
SET 05	Vegetable	18	16	88,89%	High
SET 05	Hairy	18	3	16,67%	Low
SET 05	Reward	18	8	44,44%	Medium
SET 05	Brain	18	18	100,00%	High
SET 06	Bother	15	6	40,00%	Low
SET 06	File	15	10	66,67%	Medium
SET 06	Island	15	15	100,00%	High
SET 06	Select	15	10	66,67%	Medium
SET 06	Pair	15	10	66,67%	Medium
SET 06	Angle	15	9	60,00%	Medium
SET 06	Reptile	15	13	86,67%	High
SET 06	Jaw	15	1	6,67%	Low
SET 06	Cliff	15	15	100,00%	High
SET 06	Terror	15	11	73,33%	Medium
SET 06	Head	15	2	13,33%	Low
SET 06	Walrus	15	6	40,00%	Low
SET 07	Palm	14	12	85,71%	High
SET 07	Predator	14	13	92,86%	High
SET 07	Funnel	14	6	42,86%	Low
SET 07	Refuel	14	5	35,71%	Low
SET 07	Adjustable	14	5	35,71%	Low
SET 07	Rodent	14	6	42,86%	Low
SET 07	Crash	14	3	21,43%	Low
SET 07	Flask	14	7	50,00%	Medium
SET 07	Artic	14	9	64,29%	Medium
SET 07	Calculate	14	13	92,86%	High

SET 07	Triplets	14	13	92,86%	High
SET 07	Pollute	14	13	92,86%	High
SET 08	Bouquet	14	0	0,00%	Low
SET 08	Wave	14	7	50,00%	Medium
SET 08	Get out	14	11	78,57%	Medium
SET 08	Pod	14	1	7,14%	Low
SET 08	Classify	14	13	92,86%	High
SET 08	Grapevine	14	6	42,86%	Low
SET 08	Dissect	14	8	57,14%	Medium
SET 08	Glider	14	0	0,00%	Low
SET 08	Succulent	14	5	35,71%	Low
SET 08	Pelican	14	10	71,43%	Medium
SET 08	Yacht	14	8	57,14%	Medium
SET 08	Welcome	14	14	100,00%	High
SET 09	Archer	13	13	100,00%	High
SET 09	Mammal	13	12	92,31%	High
SET 09	Composer	13	11	84,62%	High
SET 09	Oasis	13	5	38,46%	Low
SET 09	Citric	13	4	30,77%	Low
SET 09	Oil	13	12	92,31%	High
SET 09	Speedometer	13	3	23,08%	Low
SET 09	Potion	13	5	38,46%	Low
SET 09	Run up	13	6	46,15%	Medium
SET 09	Reprimand	13	10	76,92%	Medium
SET 09	Porcelain	13	4	30,77%	Low
SET 09	Large	13	9	69,23%	Medium
SET 10	Handrail	12	4	33,33%	Low
SET 10	Compass	12	8	66,67%	Medium
SET 10	Instruct	12	11	91,67%	High
SET 10	Lacking	12	5	41,67%	Low
SET 10	Infinity	12	11	91,67%	High
SET 10	Choreographic	12	8	66,67%	Medium
SET 10	Confidential	12	5	41,67%	Low
SET 10	Wedge	12	3	25,00%	Low
SET 10	Equid	12	6	50,00%	Medium
SET 10	Valve	12	3	25,00%	Low
SET 10	Harvest	12	2	16,67%	Low
SET 10	Moan	12	1	8,33%	Low

Table 6: Distribution of words in English

Source: own source

Table 7

Distribution of words in each set.

Words	Peripheral	Medium	Core
SET 01	1	6	5
SET 02	1	6	5
SET 03	2	5	5
SET 04	2	5	5
SET 05	6	2	4
SET 06	4	5	3
SET 07	5	2	5
SET 08	5	5	2
SET 09	5	3	4
SET 10	7	3	2
Total	38	42	40

Table 7: Distribution of words in each set

Source: own source

Table 8

Percentages of correct answers in English and Spanish.

Words	% English correct	% Spanish correct
SET 01	74,12%	
SET 02	74,56%	
SET 03	71,05%	
SET 04	65,74%	
SET 05	50,93%	
SET 06	60,00%	92,54%
Angle/Ángulo	60,00%	78,95%
Bother/Molestar	40,00%	89,47%
Cliff/Acantilado	100,00%	100,00%
File/Lima	66,67%	100,00%
Head/Dirigir	13,33%	100,00%
Island/Isla	100,00%	100,00%
Jaw/Mandíbula	6,67%	89,47%
Pair/Par	66,67%	94,74%
Reptile/Reptil	86,67%	89,47%
Select/Seleccionar	66,67%	89,47%
Terror/Terror	73,33%	94,74%

Walrus/Morsa	40,00%	84,21%
SET 07	62,50%	78,95%
Adjustable/Ajustable	35,71%	89,47%
Artic/Ártico	64,29%	73,68%
Calculate/Calcular	92,86%	100,00%
Crash/Colisionar	21,43%	5,26%
Flask/Termo	50,00%	73,68%
Funnel/Embudo	42,86%	68,42%
Palm/Palmera	85,71%	100,00%
Pollute/Contaminar	92,86%	100,00%
Predator/Depredador	92,86%	100,00%
Refuel/Repostar	35,71%	63,16%
Rodent/Roedor	42,86%	73,68%
Triples/Trillizos	92,86%	100,00%
SET 08	49,40%	61,84%
Bouquet/Ramo	0,00%	57,89%
Classify/Clasificar	92,86%	84,21%
Dissect/Disseccionar	57,14%	52,63%
Get out/Salir	78,57%	94,74%
Glider/Planeador	0,00%	26,32%
Grapevine/Parra	42,86%	57,89%
Pelican/Pelícano	71,43%	78,95%
Pod/Vaina	7,14%	26,32%
Succulent/Suculento	35,71%	42,11%
Wave/Oleaje	50,00%	78,95%
Welcome/Acoger	100,00%	68,42%
Yacht/Yate	57,14%	73,68%
SET 09	60,26%	60,09%
Archer/Arquero	100,00%	100,00%
Citric/Cítrico	30,77%	47,37%
Composer/Compositor	84,62%	94,74%
Large/Cuatioso	69,23%	52,63%
Mammal/Mamífero	92,31%	84,21%
Oasis/Oasis	38,46%	47,37%
Oil/Lubricar	92,31%	26,32%
Porcelain/Porcelana	30,77%	47,37%
Potion/Brebaje	38,46%	57,89%
Reprimand/Reprimenda	76,92%	63,16%
Run up/Izar	46,15%	52,63%
Speedometer/Velocímetro	23,08%	47,37%
SET 10	46,53%	57,89%
Choreographic/Coreográfico	66,67%	63,16%
Compass/Brújula	66,67%	84,21%
Confidential/Confidencial	41,67%	31,58%

Equid/Équido	50,00%	36,84%
Handrail/Barandilla	33,33%	100,00%
Harvest/Cosechar	16,67%	57,89%
Infinity/Infinito	91,67%	94,74%
Instruct/Instruir	91,67%	68,42%
Lacking/Carente	41,67%	57,89%
Moan/Gemir	8,33%	68,42%
Valve/Válvula	25,00%	15,79%
Wedge/Cuña	25,00%	15,79%
SET 11		41,20%
Total	61,51%	65,42%

Table 8: Percentages of correct answers in English and Spanish

Source: own source