Second Language Acquisition

The Influence of CLIL on Receptive Vocabulary:
A Preliminary Study

Kevin Iglesias Diéguez

Degree in English Studies

Acknowledgments

Abstract

In the last two or three decades, being able to communicate in a foreign language has become an essential trait of any European citizen due to globalisation and migration, resulting in a multi-ethnic and multilingual society. With this in mind, the European Union has been promoting the implementation of a new type of instruction that seeks to improve people's ability to communicate, namely Content and Language Integrated Learning (CLIL).

Previous studies have shown that CLIL seems to be beneficial to receptive vocabulary, which in turn correlates with a higher level of general competence. However, these studies have mainly compared CLIL and Non-CLIL groups matching in the year of instruction, which means that other factors could explain the variation found. The present study, even though exploratory in nature, sets out to fill this gap by comparing groups with the same onset age as well as controlling for other variables, such as the number of hours of exposure. This way, any improvement, or lack thereof, can be traced to the type of instruction.

The sample consisted in students from 1st and 3rd year of Compulsory Secondary Education (known as *ESO* for its Spanish name) who had started learning English at the age of 3. They were divided into groups depending on whether they were taught any subject through English (in addition to English lessons) and what grade they were in. Other sources of exposure to English have been carefully controlled for. To test general proficiency, the Quick Placement Test (QPT) was used, and the 1,000 and 2,000 frequency bands of the Vocabulary Levels Test (VLT) were delivered to measure functional vocabulary size.

Results show that vocabulary forms an integral part of general proficiency and suggest that its relevance increases as level of mastery in the target language improves. Moreover, CLIL students in 1st and 3rd ESO have outstripped their respective Non-CLIL counterparts in both general proficiency and receptive lexical knowledge, which means they have attained a better ability to understand a foreign language with the same years of instruction. In addition, CLIL learners in 1st ESO have been found to perform as well as a Non-CLIL sample in 3rd ESO with 57 more hours of exposure in general

proficiency and functional vocabulary size. Considering the level of English language lessons, differences in cognitive maturity and lower amount of exposure, it is argued that CLIL instruction has benefits beyond allowing more hours of English instruction in the same number of academic years of study. The present paper suggests that CLIL implementation should be further encouraged.

Keywords: Content and Language Integrated Learning (CLIL), receptive vocabulary, general proficiency, L3 English

Table of Contents

1. Introduction

Over the last two or three decades, research in Second Language Acquisition (SLA) has experienced a boom in two relatively-new areas: vocabulary knowledge as an important part of linguistic competence and Content and Language Integrated Learning (CLIL) as a new type of instruction of a foreign language.

Although lexical competence does not guarantee high communicative proficiency, it is a fundamental pillar of language use, which will in turn facilitate communication (Nation, 1993; Nation & Waring, 1997; Meara, 1996). The oft-cited image of acquirers carrying dictionaries instead of grammar books (Krashen, 1989) is far from being mere anecdotal evidence. Empirical research has found strong positive correlations between vocabulary and the so-called "passive skills" of reading and listening. Laufer (1992) shows strong positive correlations between two vocabulary tests and a reading test. Being part of a large project called DIALANG, Alderson (2005) reports similar coefficients between different aspects of lexical competence and reading and listening comprehension, which are in accordance with Qian (1999) and Nemati (2010). Although correlations do not point to a cause-effect relationship, it is important to point out that there is always a rather strong correlation between vocabulary and passive comprehension.

Regarding CLIL, a vast amount of research has been devoted to clarifying whether the type of instruction has any effect on the acquisition of a Foreign Language (FL). As Ellis (1994: 17) points out, this line of research has been motivated by "a desire to address issues of general theoretical interest to SLA research and also by a desire to improve the efficacy of language pedagogy." Whereas the main focus has been traditionally given to the study of grammar acquisition, it has been only recently that attention is being paid to the effect of CLIL on vocabulary learning, arguably because of its importance in achieving communication.

However, much of the research has centred only on giving a measurement of the receptive vocabulary size of L2 students. For this reason, studies dealing with the effect

of CLIL over traditional teaching methods are still scarce, especially in secondary education. Such studies have mainly compared CLIL and Non-CLIL groups matching in the year of instruction, which means that other factors could explain the variation found. The present paper will contribute to filling this gap by carrying out a pseudo-longitudinal study in which I will compare the size of receptive vocabulary of learners in 1st and 3rd year of Compulsory Secondary Education[1] in two different instructional contexts, namely CLIL and English as a Foreign Language (EFL) teaching, while controlling for several variables such as onset age and the number of hours of exposure.

To this end, the paper is organised in the following sections. Firstly, a brief description of what CLIL is and how it has been put into practice in Europe and in the Basque Country will be given. Then, the use of the Vocabulary Levels Test (VLT) as an instrument to measure receptive vocabulary will be justified. After this, I will review previous studies that have been carried out in the field of vocabulary acquisition both in CLIL and Non-CLIL contexts, which will serve as a comparison with the results of the present study. Finally, the results will be reported together with a discussion of their possible causes and implications.

2. CLIL

Since the mid-90's, there has been a growing concern in the European Union regarding people's ability to communicate in a language that is not their mother tongue. As Ruiz de Zarobe (2008) notes, this interest arises from the need to create a more inclusive and integrative society, mainly as a way to cope with a multiethnic reality. However, the European Commission's (June, 2012) report shows that most people consider learning a new language beneficial for work or study-related prospects since these answers occupy 4 out of the 5 most chosen options, with "understanding people from other countries" ranking sixth. The pursuit of different objectives, together with little guidance from European institutions (Ruiz de Zarobe, 2013), has caused the CLIL type of instruction to be implemented in different ways depending not only on the country, but also on the region and individual ventures, as well as other contextual factors that "influence both their aims and outcomes" (Nikula, Dalton-Puffer &

---

[1] In Spain, Compulsory Secondary Education is known as *Educación Secundaria Obligatoria* (ESO), and post-secondary education is referred to as *Bachillerato*. Throughout this paper, the terms *ESO* and *Bachillerato* will be used for these educational stages.

Llinares, 2013: 72). As a consequence, CLIL is often used as an umbrella term to describe any approach where "a second language is used to teach certain content subjects in the curriculum other than language lessons" (Canga, 2015)[2].

In the Basque Country, CLIL has been recently implemented on top of the existing 3-model system for bilingual education. Model A offers all subjects in Spanish, Model B teaches half the subjects through Spanish and the other half through Basque, and Model D is entirely carried out in Basque, with the exception of language lessons in all cases. The Basque Government is currently encouraging schools' adhesion to the *Framework for Trilingual Education* (Marco de Educación Trilingüe, MET), which is a policy that aims to make learners proficient in three languages, namely Spanish, Basque and English. However, schools differ on several factors which make CLIL implementation rather heterogeneous: starting age for CLIL instruction, number of subjects and hours offered in English, and electiveness of such subjects, among others.

Generally, CLIL pedagogy is characterised by a more student-centred approach, as opposed to the traditional teacher-centred one, with the focus on students' participation and interactions using the target language in an attempt to develop their communicative competence. This is achieved by providing them with comprehensible input in addition to a more "natural" context for acquisition and encouraging interaction on the students' part. However, the culture of the CLIL classroom is still that of the L1, and as Ruiz de Zarobe (2013: 237) notes, "the teachers' pragmatic use of the language is sometimes less varied than in the teaching of subjects in the L1," depending on teacher's proficiency in the target language.

Moreover, teachers of CLIL subjects are not language teachers and concentrate mainly on content rather than form (Navés, 2009). For this reason, research suggests that whilst general proficiency is improved, specific aspects of language do not seem to behave in the same way. As for general proficiency, Ruiz de Zarobe (2008) compared the performance of CLIL and Non-CLIL students at the end of Secondary and Post-

---

[2] The term *Content Based Instruction* (CBI), among others, is also used for the kind of instruction described here. Although each term is associated with its historical genesis, their actual current pedagogies do not differ to such an extent so as to consider them different (Cenoz, Genesee & Gorter, 2014; Cenoz, 2015). In consequence, any aspect other than teaching content subjects in English should be pointed out explicitly if they are considered to have an effect on the feature(s) under study, such as exposure to the target language outside the classroom.

Secondary Education in written and oral production. She reports significant differences in the four aspects of general proficiency tested in favour of CLIL learners. Lasagabaster (2008) found that CLIL students outperformed a sample of Non-CLIL learners matching in number of hours of exposure and another sample matching in grade in terms of speaking, listening, grammar and writing tasks. However, benefits of CLIL do not seem to extend to some specific areas of language such as morphosyntax. García Mayo and Villarreal Olaizola (2010) report no significant differences between students in a CLIL setting and learners in traditional instruction in different tense and agreement morphemes. Similarly, Martínez Adrián and Gutiérrez Mangado (2015) analysed the general proficiency and the production of tense and agreement morphemes in a group of CLIL students, a group of Non-CLIL students matching in year of instruction and hours of exposure, and another group with the same number of hours of exposure but different age at testing. They found that CLIL learners could perform as well as older students and that they outstripped students of the same age but with fewer hours of exposure when tested for general proficiency. Regarding specific aspects of morphosyntax, CLIL students are reported to have obtained similar results to the group with fewer hours of exposure and significantly poorer scores than older learners. In order to solve difficulties found in these aspects, focus-on-form has been proposed by several researchers (García Mayo, 2012; Ruiz de Zarobe & Lasagabaster, 2010).

It still remains to be seen whether the trend found in specific areas of grammar also occurs in receptive vocabulary. To this end, the following sections of the paper will focus on lexical knowledge, starting with the justification of an instrument that measures this construct.

3. The VLT as a Measure of Receptive Vocabulary

In this section, I will highlight the different aspects that legitimise the VLT as a measuring instrument of receptive vocabulary size.

Firstly, corpus frequency is used as the benchmark of vocabulary size in the VLT. It is assumed that the more frequent a word is, the more useful this word will prove to be. The rationale behind this assumption is as follows:

> If an item naturally occurs frequently in the language being taught, it is likely to be important also for the target behaviour of the learner: the learner will later often come across that item in reading and listening [...]. (Leech, 2001: 1)

Thus, the VLT attempts to describe the set of total words known that are functionally important, which form the learners' vocabulary profile. In order for this principle to be valid, we need to assume that learners acquire vocabulary according to its frequency. This is supported by an observed decline in students' scores as they move from high to lower frequency bands (Read, 1988; Laufer, Elder, Hill, & Congdon, 2004).

Regarding the number of words known, two different counting systems have been used in the literature. The first one uses lemmas as the basic unit for receptive vocabulary, which consist of "a headword and some of its inflected and reduced forms," such as contractions (Nation, 2001: 7). This approach has the advantage of including in the same category different tokens that only differ in grammatical information but it does not take into account transparently derived words. The second method overcomes this problem by incorporating in the same category words that are formed with affixes ranking up to Level 5 in Bauer and Nation's (1993) hierarchy. These categories are referred to as word-families (Nation, 2001) and they are those that the VLT is built upon.

However, this newer method is still far from perfect. As Bogaards (2001) points out, a word-family approach fails to account for polysemy since it presupposes that a learner that knows the meaning of *bank* in (a) will also know its meaning in (b):

(a) He went to the <u>bank</u> to deposit some money.
(b) The river burst its <u>banks</u>.

Still, the definitions asked for in the VLT correspond to the first meaning found in the dictionary entry, which is the most widely spread meaning.

Moreover, Gyllstad (2013) notices that lexical items larger than one orthographic word (known as 'formulaic language') are ubiquitous. Estimates range from 32% (Foster, 2001) to 58% (Erman & Warren, 2000) of L1 language use. This poses a problem to the measurement of receptive vocabulary since word-families do not include multi-word units. Nonetheless, it is likely that formulaic language affects productive

vocabulary to a greater extent than it does receptive vocabulary because the meaning of some expressions that fall under this term can be transparently derived since they comply with the compositionality of meaning to all effects and purposes, as in *(make) an informed decision*. Although it can sometimes be incomplete, the word-families approach used in the VLT seems to be useful in estimating vocabulary size.

Another aspect that legitimises the VLT as a measuring instrument is the visual organisation of the items: the layout of this test avoids guessing, which allows for a more accurate estimate of vocabulary size, and favours consistency. This is achieved by distributing words at each level in clusters according to their morphological category (in the same proportion as observed in the corpus)[3] and inserting distractors which bear no resemblance to the definitions given (as far as possible). In addition, the items in each cluster are alphabetically ordered and the selected ones are chosen randomly, so that no pattern emerges throughout the test. As a result, each VLT is composed of 10 clusters like the following:

```
1 adopt
2 climb       _____ go up
3 examine     _____ look at closely
4 pour        _____ be on every side
5 satisfy
6 surround
```

Finally, the validity and reliability of the VLT have been verified by empirical research. Schmitt, Schmitt and Clapham (2001) reported that the vocabulary profile targeted by the VLT falls into an implicational scale since students' scores are highly scalable. They also found that this test reflects actual lexical knowledge by carrying out interviews with the participants. Moreover, a factor analysis showed that the test is essentially unidimensional. Similarly, Beglar and Hunt's (1999) study indicates that the VLT is acceptably reliable, which lies in accordance with other studies (Xing & Fulcher, 2007).

All in all, the VLT provides a quite accurate yet not comprehensive measure of receptive vocabulary size. Although there have been attempts to devise a list of formulaic language (González & Schmitt, 2015; Martínez & Schmitt, 2012), there has

---

[3] The proportion of word classes in the English corpus used for the elaboration of the VLT falls into a 3 (noun) : 2 (verb) : 1 (adjective) ratio.

not been any breakthrough in overcoming the methodological difficulties that arise when trying to combine word families and formulaic language to measure functional vocabulary knowledge. To date, the VLT remains amongst the most valid and reliable instruments available for such purposes. As a result, it is widely used in vocabulary studies, of which the most relevant for this paper will be reviewed in the next section.

4. Studies on Receptive Vocabulary

In this section, I will review different studies that have been carried out on receptive vocabulary both in CLIL and Non-CLIL groups.

Lexical research has been conducted along three different lines: the size of receptive or productive vocabulary and how much it increases after a period of time, the relationship between these two constructs with one another as well as with general proficiency, and how a given variable influences vocabulary size.

Regarding the first case, these studies are more concerned with the mastering of vocabulary in order to improve certain areas of specific proficiency, that is, they aim at checking the efficiency of teaching in schools and the level students have in lexical knowledge (Canga, 2013). As for the second case, research is more centred on the acquisition aspect, where a model of vocabulary acquisition is being sought (De Bot, Sima Paribakht, & Bingham Wesche, 1997). Finally, variable-focused studies pursue a similar objective to pure longitudinal (or pseudo-longitudinal) studies since they set out to ascertain which factors yield better results. However, they also share some goals with those studies that delve into the relationship between different aspects of proficiency in that the study of a variable can give some guidelines as to how the brain acquires vocabulary. Moreover, research in the conditions in which vocabulary acquisition takes place can shed light on the ever-lasting debate on whether the acquisition of a first language (L1) differs from that of a second (L2) or additional (Ln) language. As multidisciplinary as they are, this last type of studies has been conducted only over the last decade or so.

The variables under inspection have generally centred on the level of motivation; age at which first exposure took place; age at the time of testing; the effects of maturity

and memory; and the type of instruction, which revolves around the influence of CLIL on vocabulary over traditional EFL teaching. This approach has gained special importance in the last decade since CLIL projects are increasingly being implemented in schools throughout Europe. Nonetheless, they are still very scarce and most comparisons need to rely on previously conducted studies which aimed at finding vocabulary size estimates, with subjects having received different hours of exposure (both inside and outside the classroom) and having started learning English at different ages.

In this respect, studies in Non-CLIL contexts abound, especially in the last years of Primary Education. Jiménez Catalán and Terrazas (2005-2008) report 4th graders' receptive vocabulary to be around 737 words after 419 hours of instruction in English. Terrazas and Llach (2009) found a much lower estimate of 361 words for 4th graders after the same amount of exposure. In the same study, the size for students' vocabulary in 5th, 6th Primary and 1st ESO was calculated at 509, 631 and 817 words, respectively. Llach and Terrazas (2012) conducted a cross-sectional study in all the grades between 4th Primary and 3rd ESO. They report similar results to those in Gallego and Llach (2009), while the estimates for 2nd and 3rd ESO are 987 and 1206 words. Canga (2013) analysed lexical knowledge of 4th ESO students and results show a mean of 935 words after 1049 hours of instruction, which is a poorer score than 2nd and 3rd ESO students' in Llach and Terrazas (2012). In an in-depth investigation about vocabulary tests, López-Mezquita (2005) carried out a not-so-controlled study of students in 4th ESO and 1st and 2nd Bachillerato. As for the first group, results point to a knowledge of 941 words, similar to Canga's (2013) but considerably lower than Llach and Terrazas (2012). This may be due to the inclusion of students who had failed the subject of English language in previous years but had passed on to the next educational level nonetheless, as well as learners who were in curricular diversification programs with much lower standards than typical 4th ESO groups. It should be noted that no variable was controlled for with the exception of grade. In similar circumstances, estimates for 1st and 2nd Bachillerato amount to 1582 and 1885 words, respectively.

As far as the CLIL variable is concerned, studies on lexical knowledge are scarce mainly due to the difficulty of finding homogeneous groups of subjects. Jiménez Catalán and Ruiz de Zarobe (2009) compared 6th grade students in CLIL and Non-CLIL

contexts, with the latter receiving 331 hours of exposure less than the former. Moreover, the CLIL group was composed of only bilinguals whereas the latter were exclusively monolinguals. For the CLIL group, the size of their vocabulary is estimated at 748 words, and for the Non-CLIL, at 602 words. In a similar study, Fernández Fontecha (2014a) set up a group of CLIL students in 5th Primary and another one of Non-CLIL learners in 2nd ESO, both of whom had received approximately 839 hours of instruction. Results show that the Non-CLIL group outperformed the CLIL group: 985 words for the former and 705 for the latter. Canga (2015) compared the scores of three groups: two 6th-grade samples (one with CLIL and the other one with traditional teaching) and one formed by 2nd ESO students in a traditional Non-CLIL context. He reports slightly higher scores for the secondary Non-CLIL group, although statistical analysis shows that this difference is not significant. A summary of the results, together with the variables considered, is shown in Table 1.

Study	Year/Grade	CLIL?	Hours of exposure	Vocabulary size
Jiménez Catalán and Terrazas (2005-2008)	4th Primary	N	419	737
Terrazas and Llach (2009)	4th Primary	N	419	361
	5th Primary	N	524	509
	6th Primary	N	629	631
	1st ESO	N	734	817
Llach and Terrazas (2012)	4th Primary	N	419	361
	5th Primary	N	524	527
	6th Primary	N	629	663
	1st ESO	N	734	836
	2nd ESO	N	839	987
	3rd ESO	N	944	1206
Canga (2013)	4th ESO	N	1049	935
López-Mezquita (2005)	4th ESO	N	-	941
	1st Bachillerato	N	-	1582
	2nd Bachillerato	N	-	1885
Jiménez Catalán and Ruiz de Zarobe (2009)	6th Primary	N	629	602
	6th Primary	Y	960	748
Fernández Fontecha (2014a)	5th Primary	Y	839	705
	2nd ESO	N	839	985
Canga (2015)	6th Primary	N	629	601
	6th Primary	Y	944	903
	4th ESO	N	1049	936

Table 1. Summary of word estimates for different grades in vocabulary research.

As can be seen from López-Mezquita (2005), caution should be taken since insightful comparisons and estimations can only be made when the variables are controlled for; otherwise, results will vary wildly. Apart from the study conducted by López-Mezquita (2005), the rest of the investigations are better-designed in this respect but are nonetheless limited in that they do not take into consideration the rate of acquisition rather than the end result. As a consequence, there is no paper which sheds light on whether the CLIL type of instruction actually has a beneficial impact on vocabulary size in addition to allowing a greater number of hours of exposure. The present study purports to fill this gap by comparing CLIL and Non-CLIL subjects who have started learning English at the same age, have received similar amounts of exposure and have not taken any extracurricular activity in English (cf. section 6.1).

5. Research Questions

The present paper aims at overcoming the limitations of previous studies and pointing towards a clearer answer to those questions that have been raised in the literature. Firstly, I checked whether the sample used in this study shows a correlation between receptive vocabulary and general proficiency, as has been previously reported (Nemati, 2010; Qian, 1999). In addition to showing the relationship between these two aspects, it also serves the purpose of assessing the representativeness of the sample. Secondly, the results of CLIL and Non-CLIL groups with the same age at testing were compared to ascertain the academic success and effectiveness of CLIL in receptive vocabulary. Since previous studies have mainly focused on the end result (Canga, 2015; Jiménez Catalán & Ruiz de Zarobe, 2009), this paper includes a pseudo-longitudinal analysis to reach more insightful conclusions. Finally, I compared a CLIL and a Non-CLIL group with the same onset age and number of hours of exposure but a difference in testing age to assess the impact of CLIL on receptive lexical knowledge, overcoming the limitations found in Fernández Fontecha (2014a), and similar in design to Lasagabaster's (2008) assessment of the relationship between CLIL and general proficiency. In short, this study addresses the following questions:

1. Is there a relationship between receptive vocabulary and general proficiency?
2. Do students in a CLIL context outperform their counterparts in a traditional EFL classroom at the same educational level?
3. Do CLIL students outperform older Non-CLIL students when they have been exposed to the same number of hours and all other factors are held constant?

6. The Study

6.1. Methodology

I. PARTICIPANTS

The sample consisted of 55 Basque-Spanish bilingual students from three different schools learning English as a L3. Participants come from both Basque-speaking and Spanish-speaking families, but are to be considered equally competent in both languages since Basque is the language of instruction (Cenoz & Valencia, 1994). They all share a similar socio-economic background since the three schools that participated in the study are located in Portugalete (on the left-margin of the river Nervión), where socio-economic status is rather homogeneous throughout. None of the subjects attended an academy or any extracurricular activities related to English, nor had they made any trip to an English-speaking country.

Participants were divided into four groups considering their type of instruction and their current year of instruction, which determines the number of hours of exposure[4]: (a) a Non-CLIL 1 group (n=10) of 12 year-olds in 1st ESO; (b) a CLIL 1 group (n=15) with the same age as the previous group but more hours of exposure; (c) a Non-CLIL 2 group (n=15) of 14 year-olds in 3rd ESO with a similar number of hours of exposure to the CLIL 1 group; and (d) a CLIL 2 group (n=15) with students of the same age, also in 3rd ESO. Only participants who started learning English at the age of 3 have been included in the sample. In doing so, this study purports to overcome the limitations that have arisen in previous studies dealing with the effect of CLIL instruction on receptive vocabulary due to the lack of matching between the number of hours of exposure and the onset age. Participants' characteristics are displayed in Table 2.

Group	Onset Age	Age at testing	Length of exposure (in years)	Hours of Exposure
Non-CLIL 1 (n=10)	3	12	9	972
CLIL 1 (n=15)	3	12	9	1,116
Non-CLIL 2 (n=15)	3	14	11	1,173
CLIL 2 (n=15)	3	14	11	1,451

Table 2. Participants' characteristics.

---

[4] The variable gender has not been taken into consideration when dividing the participants into groups since differences between males' and females' learning behaviours in lexical learning seem to be test-dependent (Sunderland, 2010), with only small differences arising at some stages due to psychological changes characteristic of puberty and motivational factors related to adolescence (Llach & Terrazas, 2012).

Non-CLIL groups had 2 hours of English a week during their first three years of formal education and 3 hours a week of EFL lessons in Primary and Secondary Education. In addition to these hours of formal instruction of English, the CLIL groups had 1 hour a week of CLIL in Social Sciences and Creative Arts during the last 3 years of Primary Education and throughout Secondary Education. As a result, the CLIL 1 group had a total of 4 years of CLIL instruction, whereas the CLIL 2 group were exposed to CLIL for 6 years. In all groups, the materials and approach used in the English lessons were the same, with the occasional use of Spanish or Basque when needed.

II. INSTRUMENTS

Data were gathered by means of three instruments. To measure the participants' general proficiency, the first part of the QPT (version 1) was used (Appendix 2). Part 2 was not handed out since it corresponds to proficiency levels of mastery[5], which are beyond the scope of Secondary Education. This test has been extensively used in SLA research to assess general proficiency (López-Mezquita, 2005; Martínez Adrián & Gutiérrez Mangado, 2015). In addition, two different VLTs were used, namely the 1,000 and 2,000 frequency bands, to measure the size of students' receptive vocabulary (Appendices 3 and 4). The 1,000 VLT consisted in translation of words to avoid difficulties arising from not understanding the definitions rather than the target vocabulary items[6]. The 2,000 VLT is a slightly modified version of the test developed by Schmitt, Schmitt and Clapham (2001) that has been previously used in Fernández Fontecha (2014a). Tests for both frequency bands were used with the purpose of getting an insightful understanding of students' functional vocabulary regarding the 1,000 and 2,000 most frequent words, since scoring at least 15 (max=30) in the 2,000 VLT is claimed to show that students master the whole 1,000 most frequent words, whilst this may not always be the case. Although students are reported to learn the most frequent words first (Read, 1988), they may also possess some significant word knowledge pertaining to the 2,000 frequency band while not mastering the previous 1,000[7]. These

---

[5] C1 and C2 levels according to the CEFR scale (https://www.coe.int/t/dg4/linguistic/Source/Framework_ EN.pdf).

[6] It has been translated by the GLAUR research group of La Rioja under the supervision and approval of Paul Nation.

[7] Lower frequency bands (5,000 and 10,000) have not been used since less frequent words hardly ever appear in Secondary Education textbooks or class materials, including CLIL, where difficulty of vocabulary is kept to a minimum to facilitate content learning, according to teachers.

tests have been empirically proved to be reliable and valid as a measure of the intended functional vocabulary (cf. section 3) and are widely used in vocabulary research (Jiménez & Terrazas, 2005-2008; Qian, 2002; Terrazas & Llach, 2009).

III. PROCEDURE

All tests were done in one session during class time, except for the questionnaire, which students were asked to complete at home with their parents. They were told that the results of these tests would not in any way affect their marks in English or any other subject, and were also told to miss out any item to which they did not know the answer. For each test, they were given clear instructions, together with an example, both in written form and orally in Spanish to clarify what they were being asked to do. In the case of the QPT, the example was made up since the test does not include one.

They were first given the QPT, to be completed in 30 minutes, followed by the VLTs in order of frequency band, for which the time allotted was 10 minutes each.

Total scores and vocabulary size estimates were obtained. To this end, Nation's (1990: 78) formula was applied. Individual data were entered into SPSS for descriptive and inferential statistical analysis. The results of the Kolmogorov-Smirnov tests revealed that all groups were normally distributed in all tests. Since they complied with the normality assumption, Pearson's correlation coefficient was calculated between the QPT and the VLTs. Independent samples t-test was also implemented to check for any significant differences between the groups' means. Since Levene's test did not show to be significant, the homogeneous-variances values are reported.

6.2. Results and Discussion

I. RESEARCH QUESTION I

The first research question aims to ascertain whether a relationship exists between receptive vocabulary and general proficiency. The results for the correlations are shown in Table 3 below.

		QPT	1,000 VLT	2,000 VLT
Non-CLIL 1	**QPT**	1		
	1,000 VLT	0.198	1	
	2,000 VLT	.811**	.279	1
CLIL 1	**QPT**	1		
	1,000 VLT	.659**	1	
	2,000 VLT	.856**	.666**	1
Non-CLIL 2	**QPT**	1		
	1,000 VLT	.520*	1	
	2,000 VLT	.731**	.758**	1
CLIL 2	**QPT**	1		
	1,000 VLT	.562*	1	
	2,000 VLT	.762**	.49	1

* $p < .05$
** $p < .01$

Table 3. Pearson Correlation Coefficients for the QPT and VLTs by group (two-tailed).

In all groups, the 2,000 VLT strongly correlates with the QPT at a significant level, which suggests that vocabulary plays an essential role in proficiency. These figures (.731-.856) agree with previous studies that have purported to answer this question (Nemati, 2010; Qian, 1999).

In the case of the present study, a correlation was expected to a certain extent, since some of the questions in the QPT tap specifically on vocabulary knowledge. However, such strong coefficients point to a great relevance of vocabulary for other parts of general proficiency than simply the vocabulary compartment, highlighting the importance of lexical knowledge in successful communication as proficiency in English increases.

Moreover, by including the 1,000 VLT, interesting results arise. Its correlation with the QPT is lower than that of the 2,000 VLT (coefficients ranging from .520-.659), and the relationship between these two tests is rather weak and non-significant in the Non-CLIL 1 group. Lower coefficients may suggest that knowledge of the 1,000 most frequent words does not affect general proficiency to the extent that the next frequency band does. This seems to signal that knowledge of vocabulary from different frequency bands has a differing impact on general proficiency. For this reason, caution should be taken when reporting and interpreting results in these correlations between receptive

vocabulary and general proficiency when only one test is used for lexical knowledge. By and large, knowledge of the 1,000 most frequent words seems to have a lesser effect on general proficiency than knowledge of the next 1,000.

As for the impact of CLIL, it seems that this type of instruction slightly increases the correlation between general proficiency and receptive vocabulary knowledge, suggesting that a bigger size of receptive vocabulary relates to a more successful performance in general proficiency. However, these data should be submitted to further statistical analysis to check for any significant difference in the Pearson coefficients between groups that only differ in this variable.

All in all, functional receptive vocabulary seems to be of great importance in achieving successful communication. In this respect, CLIL is argued to increase this type of lexical knowledge when compared to traditional EFL. Since the main purpose of the educational system is to allow students to communicate in a foreign language, this type of instruction should be implemented in all schools if it really improves receptive vocabulary knowledge as reported. This is analysed in the next research question.

II. RESEARCH QUESTION II

The second research question is concerned with the differences between the CLIL groups and their Non-CLIL counterparts at the same year of instruction, which means they have been exposed to the English language for different amounts of time (cf. Table 2).

Descriptive statistics for the Quick Placement Test in 1st ESO (aged 12) are shown in Table 4, together with the results of the t-test. Tables 5 and 6 respectively show the results of the corresponding analyses for the 1,000 VLT and the 2,000 VLT.

	Non-CLIL 1	CLIL 1	t
Mean	13.30	18.47	
SD	2.58	3.48	
Min	10	14	-4.004 (p=.001)
Max	18	26	
Range	8	12	

Table 4. Descriptive and inferential statistics for the QPT in 1st ESO.

15

	Non-CLIL 1	CLIL 1	t
Mean	19.40	24.13	
SD	3.86	1.81	
Min	12	21	-4.143
Max	24	27	(p=.000)
Range	12	6	

Table 5. Descriptive and inferential statistics for the 1,000 VLT in 1st ESO.

	Non-CLIL 1	CLIL 1	t
Mean	9.73	15.53	
SD	2.68	4.03	
Min	2	10	-5.520
Max	12	25	(p=.000)
Range	10	15	

Table 6. Descriptive and inferential statistics for the 2,000 VLT in 1st ESO.

The Non-CLIL 1 group got a mean score 13.30 in the QPT, which according to the score guide, corresponds to a beginner's level (A1). This is rather disturbing since higher mastery is expected after 972 hours of exposure and learning. The CLIL 1 group obtained a higher mean of 18.47, which signals that students have achieved an elementary level (A2) after 1,116 hours of exposure. Inferential statistics shows that the difference in general proficiency is significant (p=.001) in favour of CLIL learners, which agrees with previous studies (Lasagabaster, 2008; Martínez Adrián & Gutiérrez Mangado, 2015; Ruiz de Zarobe, 2008).

Regarding the VLTs, the means are also higher for CLIL students. In the 1,000 VLT, the mean for the Non-CLIL students is 19.40, which yields a vocabulary estimate of 647 words. In the 2,000, students averaged 9.73 points, which using Nation's (1990) formula gives a receptive vocabulary size of 648 words out of 2,000. This estimate is lower than those found by Llach and Terrazas (2012) and Jiménez Catalán and Ruiz de Zarobe (2009) for students with approximately the same number of hours of exposure. The proximity of both estimates suggests that participants had hardly acquired any words from the 2,000 most frequent words that do not pertain to first 1,000. As far the CLIL group is concerned, they obtained a mean score of 24.13 points in the 1,000 VLT, corresponding to a receptive vocabulary of 804 words. According to the results of the 2,000 VLT, this group has a functional vocabulary of 1,035 words since the average

score is 15.53. This means that students already master the majority of the 1,000 most frequent words and have acquired slightly over 200 of the next frequency band. Independent-samples t-tests show that the CLIL group has outperformed the Non-CLIL group at a significant level in 1st ESO (p=.000).

As for 3rd of ESO (aged 14), the results obtained are shown in Tables 7, 8 and 9, following the same order of presentation as above.

	Non-CLIL 2	CLIL 2	t
Mean	19.93	23.33	
SD	3.13	4.19	-2.520
Min	15	16	(p=.018)
Max	25	34	
Range	10	18	

Table 7. Descriptive and inferential statistics for the QPT in 3rd ESO.

	Non-CLIL 2	CLIL 2	t
Mean	23.87	25.93	
SD	3.82	1.44	-1.963
Min	18	23	(p=.060)
Max	29	29	
Range	11	6	

Table 8. Descriptive and inferential statistics for the 1,000 VLT in 3rd ESO.

	Non-CLIL 2	CLIL 2	t
Mean	17.20	20.87	
SD	4.16	2.92	-2.792
Min	9	15	(p=.009)
Max	23	27	
Range	14	12	

Table 9. Descriptive and inferential statistics for the 2,000 VLT in 3rd ESO.

With regard to general proficiency, students in the Non-CLIL group have achieved an elementary level (A2) with a mean of 19.93 points, whereas the CLIL group is in the threshold between elementary and lower intermediate (B1) with 23.33 points on average. As in the 1st ESO case, the CLIL group has performed significantly better than the Non-CLIL group (p=.018).

In the VLTs, students in the CLIL group have also obtained higher scores. The Non-CLIL group averaged knowledge of 796 words out of the 1,000 most frequent and 1,146 words from the 2,000 most frequent ones, which means that students have not fully acquired understanding of all of the 1,000 most frequent ones but have nonetheless demonstrated knowledge of 350 words belonging to the lower frequency band. The CLIL group, on the other hand, has achieved the slightly higher score of 25.93 points in the 1,000 VLT and 20.87 in the 2,000 VLT, which stand for 864 and 1,391 words, respectively. Inferential analyses show that the difference in the knowledge of the higher frequency band is not significant (although it signals a statistical tendency), and that students in the CLIL group have larger vocabularies.

In almost all cases, learners in a CLIL context have outperformed their Non-CLIL counterparts to a significant extent, with the only exception of the 1,000 VLT in the 3rd ESO groups. Although this difference in results is not significant, the CLIL 2 group knows on average over 60 words more than their Non-CLIL 2 counterparts. In this respect, the CLIL type of instruction seems to help consolidate knowledge of the 1,000 frequency band.

All in all, CLIL instruction does indeed help grow a bigger receptive vocabulary, as well as improve general proficiency. In 1st ESO there is a sharp difference of almost 400 words between their functional vocabularies. Such disparity may stem from the decontextualised use of English in traditional EFL lessons. In this type of instruction, specific vocabulary is carefully chosen in textbooks to comply with the minimum standards. Hence, it is very likely that students may not have been exposed to words that are beyond the 1,000 most frequent ones since learning words in isolation does not require the need for words that are not specifically intended to be learnt, specially taking into consideration that teachers rarely divert from the limited vocabulary presented in textbooks. On the other hand, CLIL instruction necessarily makes use of words of higher frequency bands since words in the 1,000 frequency band are not sufficient to convey the required content in CLIL subjects. Moreover, these may in turn reinforce the vocabulary to be learnt in the Englih class, since repetition of occurrence seems to play a key role in vocabulary learning (Saragi, Nation, & Meister, 1978; Webb, 2007). In other words, the most frequent and thus repeated words appear both in the traditional

English classroom (mostly decontextualised) and in the other subjects taught through English (in context).

However, a pseudo-longitudinal analysis of the results reveals that CLIL students do not make the most of the greater number of hours of exposure. In other words, they seem to have a lower rate of acquisition. In the Non-CLIL groups, students in 3rd of ESO know 149 more words among the 1,000 most frequent ones and 498 more belonging to the next frequency band in 201 hours, which respectively averages 74 and 247 words per 100 hours of exposure to the target language. In the CLIL groups, these differences amount to a total of 60 words for the lower frequency band and 356 for the next one in 335 hours, with an average of 18 and 106 words per 100 hours of exposure.

This seems to point to CLIL not only being ineffective but rather slowing down vocabulary acquisition. However, there are two reasons that may explain these results.

Firstly, the study is not entirely longitudinal, meaning that the subjects tested in the CLIL 1 group are not the same as those in the Non-CLIL 2 group. Since three schools are involved in this study, they may have contributed unequally to the conformation of each group, which may show up and magnify slight differences in teaching that possibly make this small sample not wholly representative. For more solid grounds, a longitudinal study is warranted.

Secondly, CLIL students are not as exposed to other words other than those appearing in the English class as could be expected. A qualitative analysis of end-of-degree projects (TFGs) and Master's Dissertations dealing with CLIL didactic units has revealed that the vocabulary used almost entirely coincides with the vocabulary presented in the English language textbooks for the same year of education[8] (e.g. Calvario Pérez, 2014; Lázaro Gómez, 2013). This suggests that the difference in vocabulary between CLIL and Non-CLIL students is caused by incidental learning of vocabulary used in extra material provided by the teacher and class dynamics (such as group discussions) rather than explicit learning. Considering that English classes are the same for both CLIL and Non-CLIL students, the effect of CLIL can be deemed

---

[8] This is especially noticeable in those exercises that focus on vocabulary, where the difference between words being asked for in a traditional-English-class exercise and a CLIL subject is virtually nonexistent.

remarkable given the little attention paid to increasing receptive vocabulary. Generally, it seems that CLIL subjects reinforce the vocabulary used in the English class by allowing the repetition of these words in addition to present students with a few more words that they learn incidentally.

Nonetheless, this apparent effectiveness of CLIL instruction as regards receptive vocabulary may also be attributable to the greater number of hours of exposure to the target language. It still remains to be seen whether students in Non-CLIL settings would achieve the same results if they had an additional hour of English instruction per week (and no other subject taught through English). Since the Government decides on the hours needed for each subject, this kind of study is rendered impossible. Therefore, the only approach that may succeed in unveiling the effect of type of instruction on receptive vocabulary is to find subjects that have started learning English at the same age and have been exposed to the same number of hours, notwithstanding the difference in testing age. This is the aim of the third research question.

III. RESEARCH QUESTION III

This question aims at clarifying whether CLIL instruction is beneficial for receptive vocabulary knowledge by overcoming the limitations found in previous studies. As already stated, subjects have been exposed to the English language for the same amount of time and have started learning English at the same age but differ in age at testing. In turn, any difference between the groups' performance, or lack thereof, can arguably be attributed to the type of instruction. In the present section, the CLIL 1 group will be compared with the Non-CLIL 2. All students have first been exposed to the English language at the age of 3 and, in spite of having a different age at the time of testing, have received similar hours of English instruction, either explicitly (language lessons) or through content subjects (Social Sciences and Creative Arts). More specifically, both groups share 972 hours of formal instruction in English, with the remaining difference amounting to 144 hours of CLIL subjects in the case of the CLIL 1 group, and additional 201 hours of formal instruction in English in the Non-CLIL 2 group. Descriptive statistics for these groups can be seen in Tables 4-9, whereas word estimates and relevant inferential statistics are shown in Tables 10, 11 and 12 below[9].

---

[9] Since the score of the QPT does not have a specific meaning beyond assessing proficiency level in a 6-category scale, descriptive statistics are repeated for easy reference, together with the result of the t-test.

	CLIL 1	Non-CLIL 2	t
Mean	18.47	19.93	
SD	3.48	3.13	
Min	14	15	-1.214
Max	26	25	(p=.235)
Range	12	10	

Table 10. Descriptive and inferential statistics for the QPT.

	CLIL 1	Non-CLIL 2	t
Mean	804	796	
SD	30	127	
Min	700	600	-.245
Max	900	966	(p=.808)
Range	200	366	

Table 11. Word estimates and inferential statistics for the 1,000 VLT.

	CLIL 1	Non-CLIL 2	t
Mean	1035	1,146	
SD	259	277	
Min	666	600	-1.114
Max	1666	1533	(p=.275)
Range	1000	933	

Table 12. Words estimates and inferential statistics for the 2,000 VLT.

For the three tests administered, no significant differences have been found between the two groups, which signals that CLIL does not pose any threat to lexical knowledge nor general proficiency. A more in-depth analysis and discussion of the results will suggest that the CLIL type of instruction has more benefits than simply allotting more hours to English in fewer years of academic study.

As regards general proficiency, both groups have obtained similar results, although the Non-CLIL 2 group has scored slightly higher. However, only the first part of the QPT was administered, which means that students have not been assessed on listening, writing or speaking skills. Had these tasks been included, it is likely that the CLIL group could have outperformed the Non-CLIL one since the cornerstone of CLIL instruction is participation and interaction, rather than more controlled activities. The nature of the instrument used has thus conditioned the results to some extent, as previous research has found compelling evidence that shows CLIL learners can perform

as well as or even better than traditional EFL students with the same numbers of hours of exposure (Lasagabaster, 2008; Martínez Adrián & Gutierrez Mangado, 2015; Ruiz de Zarobe, 2010).

As for receptive vocabulary within the 1,000 most frequent words, CLIL and Non-CLIL students seem to be on equal grounds since the difference is not significant. As I argue below, this could have resulted from the application of the same standards in vocabulary to the same educational level irrespective of whether students are taking more subjects in English.

In the 2,000 frequency band, the Non-CLIL 2 group has scored on average slightly higher, although CLIL students have higher minimum and maximum. These results become remarkable if we take into account that the CLIL group has been exposed to English for 57 hours less, approximately half a year of formal English instruction. The following discussion points to several advantages stemming from CLIL instruction rather than a higher number of hours of exposure in traditional EFL teaching.

CLIL methodology is described as allowing a more naturalistic and contextual learning of the language by focusing on content, much like required if learning happened in an English-speaking country. However, research initially suggested that context does not play a role in vocabulary learning. Seibert's (1930) longitudinal study, with tests delivered after one hour, two, ten and forty days found that students learning word pairs consistently outperformed students working with words in context. In the same vein, Gershman (1970) reports a non-significant difference between word pair and contextual learning. These empirical studies used a very narrow account of context, providing words in a sentence or with a drawing. More recent research uses context in much broader terms, encompassing and "simulating" CLIL situations. Coady (1997: 286) carried out a survey of previous research and reached the conclusion that "if the language is authentic, rich in content, enjoyable, and, above all, comprehensible, then learning is more successful," which coincides with the CLIL environment in subject matter lessons. More research seems to support this claim (Nagy, 1995; Webb, 2008).

Lexical knowledge also seems to be bolstered by incidental learning, since attention is drawn to content rather than to vocabulary items (Vidal, 2011). The results of the present study show that CLIL learners have acquired almost the same number of words as the Non-CLIL students have even though the former have received 56 hours less of exposure than the Non-CLIL group. Considering that traditional teaching draws their attention to vocabulary items of higher frequency bands (since less frequent words are presented in subsequent years of academic study), I gather that CLIL learners could have performed better if vocabulary standards in English lessons were raised and lexical complexity in CLIL subjects were not so limited (or at least increased steadily). Furthermore, it would possibly be beneficial to implement incidental learning in English lessons for both CLIL and Non-CLIL students.

In addition to providing a real context for English learning and use, CLIL has been shown to increase students' motivation levels, which in turn facilitate vocabulary acquisition (Fernández Fontecha, 2014b). However, the actual extent to which instruction affects motivation highly depends on the subject matter: whereas studies in Physical Education have found no significant differences between CLIL and Non-CLIL students (Heras & Lasagabaster, 2015), research in other subjects report a significantly higher level of motivation in CLIL students (Lasagabaster, 2011; Lasagabaster & Sierra, 2009). The roots of this motivational growth are related to the meaningful and teleological use of language, since students in a traditional EFL setting find some exercises boring, unrealistic and non-significant (Lasagabaster & Sierra, 2009: 13). Hence, CLIL learners are more prone to acquiring and increasing their vocabulary.

All in all, CLIL seems to be a promising methodology for the growth of functional receptive vocabulary. Results reported here are remarkable since CLIL learners have performed as well as Non-CLIL students, who have received slightly more hours of exposure and have developed greater cognitive maturity.

7. Conclusion and Pedagogical Implications

With CLIL experiencing a boom in the last decade and communication in a FL becoming essential, the present study set out to clarify the benefits of this type of

instruction over traditional EFL teaching as regards receptive vocabulary, for this plays a key role in understanding meaning as well as being significantly related to general proficiency. Comparisons between groups of different characteristics (cf. Table 2) suggest that CLIL students not only outstrip their Non-CLIL counterparts in the same year of instruction but also perform equally well when their results are compared with older learners who have been exposed to the English language for approximately the same amount of time. Although results suggest CLIL students' rate of acquisition is slower, I have proposed that this downside stems from the pseudo-longitudinal nature of the present study (rather than purely longitudinal) and the limitations of the English syllabus. Taking into account the greater number of hours that the Non-CLIL 2 students have received, their intrinsic cognitive maturity and higher complexity presented in their English lessons, I have arguably attributed to CLIL instruction methodological characteristics that favour vocabulary learning: contextual presentation of lexical items; focus on content, which allows incidental learning; and increasing motivation levels by giving language use a communicative purpose.

In this paper, I have also suggested that the benefits of CLIL as far as receptive vocabulary is concerned can be further exploited by including English lessons especially designed to meet the needs of CLIL students. These would take into consideration the further hours of exposure to the target language and the higher number of repetitions available that are needed to learn a word. In consequence, the English curriculum for CLIL students should have higher standards as far as vocabulary is concerned. In addition, CLIL materials should not be so constrained by English lessons standards in this respect and vocabulary of higher complexity should be included progressively.

Nevertheless, the results and conclusions of this paper need to be taken with caution due to the small-scale nature of the study. Had more subjects taken part, different results could have arisen. Moreover, receptive vocabulary has only been assessed on the basis of individual words, hence neglecting formulaic language. A more comprehensive estimate could have been obtained by delivering more vocabulary tests, such as the Word Associates Test, Size Test, Eurocentres Vocabulary Test or a test

based on the PHRASE and PHaVE lists[10]. It would also be interesting to test the impact of CLIL on students' academic vocabulary as well as productive vocabulary.

All in all, CLIL seems to be beneficial both in terms of general proficiency and receptive vocabulary knowledge, and since all learners could benefit from this type of instruction, its implementation should be further encouraged from educational institutions.

---

[10] Accessible at http://www.norbertschmitt.co.uk/resources.html.

References

Alderson, J. C. (2005). *Diagnosing Foreign Language Proficiency*. London: Continuum.

Bauer, L., & Nation, I. S. P. (1993). Word families. *International Journal of Lexicography, 6*, 1-27.

Beglar, D., & Hunt, A. (1999). Revising and validating the 2000 Word Level and University Word Level Vocabulary Tests. *Language Testing, 16*(2), 131-162.

Bogaards, P. (2001). Lexical units and the learning of foreign language vocabulary. *Studies in Second Language Acquisition, 23*(3), 321-343.

Calvario Pérez, J. J. (2014). *Elaboration of materials for a Conocimiento del Medio class*. Zaragoza: University of Zaragoza.

Canga, A. (2013). Receptive vocabulary size of secondary Spanish EFL learners. *Revista de Lingüística y Lenguas Aplicadas, 8*, 66-75.

Canga, A. (2015). Receptive vocabulary of CLIL and Non-CLIL Primary and Secondary school learners. *Complutense Journal of English Studies, 23*, 59-77.

Cenoz, J., Genesee, F., & Gorter, D. (2014). Critical analysis of CLIL: Taking stock and looking forward. *Applied Linguistics, 35*, 243-262.

Cenoz, J., & Valencia, J. (1994). Additive trilingualism: Evidence from the Basque Country. *Applied Psycholinguistics, 15*, 195-207.

Cenoz, J. (2015) Content-Based Instruction and Content and Language Integrated Learning: the same or different? *Language, Culture and Curriculum, 28*(1), 8-24.

Coady, J. (1997). L2 vocabulary acquisition: A synthesis of the research. In J. Coady and T. Huckin (Eds.), *Second Language Vocabulary Acquisition* (pp. 273-290). Cambridge: Cambridge University Press.

De Bot, K., Sima Paribakht, T., & Bingham Wesche, M. (1997). Toward a lexical processing model for the study of second language vocabulary acquisition. *Studies in Second Language Acquisition, 03*, 309-329.

Ellis, R. (1994). *The Study of Second Language Acquisition*. Oxford: Oxford University Press.

Erman, B. and Warren, B. (2000). The idiom principle and the open choice principle. *Text 20, 1*, 29-62.

European Commission (June, 2012). *Special Eurobarometer 386: Europeans and their languages*. Retrieved from: http://ec.europa.eu/public_opinion/archives/ebs /ebs_386_en.pdf

Fernández Fontecha, A. (2014a). Receptive vocabulary knowledge and motivation in CLIL and EFL. *Revista de Lingüística y Lenguas Aplicadas, 9*, 23-32.

Fernández Fontecha, A. (2014b). Motivation and gender effect in receptive vocabulary learning: An exploratory analysis in CLIL primary education. *Latin American Journal of Content and Language Integrated Learning, 7*(2), 27-49.

Foster, P. (2001). Rules and routines: A consideration of their role in the task-based language production of native and non-native speakers. In Bygate, M., Skehan, P., & Swain, M. (Eds), *Researching Pedagogic Tasks: Second Language Learning, Teaching, and Testing* (pp. 75-94). Harlow: Longman.

García Mayo, M. P. (2012). The relevance of attention to form in communicative classroom contexts. *Estudios de Lingüística Inglesa Aplicada, 11*, 11-45.

García Mayo, M. P., & Villarreal Olaizola, I. (2010). The development of suppletive and affixal tense and agreement morphemes in the L3 English of Basque-Spanish bilinguals. *Second Language Research, 27*(1), 129-149.

Gershman, S. J. (1970). *Foreign Language Vocabulary learning under seven conditions*. Columbia University: Doctoral thesis.

González, B., & Schmitt, N. (2015). How much collocation knowledge do L2 learners have? The effects of frequency and amount of exposure. *International Journal of Applied Linguistics, 166*(1), 94-126.

Gyllstad, H. (2013). Looking at L2 vocabulary knowledge dimensions from an assessment perspective – challenges and potential solutions. In C. Bardell, C. Lindqvist & B. Laufer (Eds.), *L2 vocabulary acquisition, knowledge and use: new perspectives on assessment and corpus analysis* (pp. 11-28). Amsterdam: EuroSLA.

Heras, A., & Lasagabaster, D. (2015). The impact of CLIL on affective factors and vocabulary learning. *Language Teaching Research, 19*(1), 70-88.

Jiménez Catalán, R. M., & Terrazas, M. (2005-2008). The receptive vocabulary of English Foreign Language young learners. *Journal of English Studies, 5-6*, 173-191.

Jiménez Catalán, R. M., & Ruiz de Zarobe, Y. (2009). The receptive vocabulary of EFL learners in two instructional contexts: CLIL versus Non-CLIL instruction. In Y. Ruiz de Zarobe & R. M. Jiménez (Eds.), *Content and Language Integrated Learning: Evidence from Research in Europe* (pp. 81-92). Bristol: Multilingual Matters.

Krashen, S. D. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *The Modern Language Journal, 73*(4), 440-464.

Lasagabaster, D. (2008). Foreign language competence in content and language integrated courses. *The Open Applied Linguistics Journal, 1*, 30-41.

Lasagabaster, D., & Sierra, J. M. (2008). Language attitudes in CLIL and traditional EFL classes. *International CLIL Research Journal, 1*(2), 4-17.

Lasagabaster, D. (2011). English achievement and student motivation in CLIL and EFL settings. *Innovation in Language Learning and Teaching, 5*(1), 3-18.

Laufer, B. (1992). How much lexis is necessary for reading comprehension? In P. J. L. Arnaud & H. Béjoint (Eds.), *Vocabulary and Applied Linguistics* (pp. 126-132). London: Macmillan.

Laufer, B., Elder, C., Hill, K., & Congdon, P. (2004). Size and strength: Do we need both to measure vocabulary knowledge? *Language Testing, 21*, 202-226.

Lázaro Gómez, M. C. (2013). *La enseñanza de contenidos en lengua inglesa: Estrategias, técnicas y recursos*. Valladolid: University of Valladolid.

Leech, G. (2001). The role of frequency in ELT: New corpus evidence brings a re-appraisal. In H. Wenzhong (Ed.), *ELT in China 2001: Papers presented at the 3rd international symposium in China* (pp. 1-23). Beijing: Foreign Language Teaching and Research Press.

Llach, M. P. A., & Terrazas, M. (2012). Vocabulary knowledge development and gender differences in a second language. *Estudios de Lingüística Inglesa Aplicada, 12*, 45-75.

López-Mezquita, M.T. (2005). *La evaluación de la competencia léxica: Tests de vocabulario. Su fiabilidad y validez*. Universidad de Granada: Doctoral dissertation.

Martínez Adrián, M., & Gutiérrez Mangado, M. J. (2015). Is CLIL instruction beneficial in terms of general proficiency and specific areas of grammar? *Journal of Immersion and Content-Based Language Education, 3*(1), 51-76.

Martínez, R., & Schmitt, N. (2012). A phrasal expressions list. *Applied Linguistics, 33*(3), 299-320.

Meara, P. (1996). The dimensions of lexical competence. In G. Brown, K. Malmkjaer & J. Williams (Eds.), *Performance and Competence in Second Language Acquisition* (pp. 35-53). Cambridge: Cambridge University Press.

Nagy, W. E. (November, 1995). On the role of context in first- and second-language vocabulary learning. *Center for the Study of Reading Technical Report, 627*. Illinois: University of Illinois at Urbana-Champaign.

Nation, I. S. P. (1990). *Teaching and Learning Vocabulary*. New York: Newbury.

Nation, I. S. P. (1993). Vocabulary size, growth and use. In R. Schreuder & B. Weltens (Eds.), *The Bilingual Lexicon* (pp. 115-134). Amsterdam/Philadelphia: John Benjamins.

Nation, I. S. P. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.

Nation, I. S. P., & Waring, R. (1997). Vocabulary size, text coverage, and word lists. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, Acquisition and Pedagogy* (pp. 6-19). Cambridge: Cambridge University Press.

Navés, T. (2009). Effective Content and Language Integrated Learning (CLIL) Programmes. In Y. Ruiz de Zarobe & R. M. Jiménez (Eds.), *Content and Language Integrated Learning: Evidence from Research in Europe* (pp. 22-40). Bristol: Multilingual Matters.

Nemati, A. (2010). Proficiency and size of receptive vocabulary: Comparing EFL and ESL environments. *International Journal of Educational Research and Technology, 1*(1), 46-53.

Nikula, T., Dalton-Puffer, C., & Llinares, A. (2013). CLIL classroom discourse: Research from Europe. *Journal of Immersion and Content-Based Language Education, 1*, 70-100.

Qian, D. D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review, 56*, 282-308.

Qian, D. D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning, 52*(3), 513-536.

Read, J. (1988). Measuring the vocabulary knowledge of second language learners. *RELC Journal, 19*(2), 12-25.

Ruiz de Zarobe, Y. (2008). CLIL and Foreign Language learning: A longitudinal study in the Basque Country. *International CLIL Research Journal, 1*(1), 60-73.

Ruiz de Zarobe, Y. (2010). Written production and CLIL: An empirical study. In C. Dalton-Puffer, T. Nikula & U. Smit (Eds.), *Language use in CLIL.* Amsterdam: John Benjamins.

Ruiz de Zarobe, Y. (2013). CLIL implementation: From policy-makers to individual initiatives. *International Journal of Bilingual Education and Bilingualism, 16*(3), 231-243.

Ruiz de Zarobe, Y., & Lasagabaster, D. (2010). CLIL in a bilingual community: The Basque autonomous region. In D. Lasagabaster & Y. Ruiz de Zarobe (Eds.), *CLIL in Spain: Implementation, results and teacher training* (pp. 12-29). Newcastle upon Tyne: Cambridge Scholars Publishing.

Saragi, T., Nation, I. S. P. & Meister, G. F. (1978). Vocabulary learning and reading. *System, 6*, 72-78.

Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing, 18*(1), 55-88.

Seibert, L. C. (1930). An experiment on the relative efficiency of studying French vocabulary in associated pairs versus studying French vocabulary in context. *Journal of Education Psychology, 21*, 297-314.

Sunderland, J. (2010). Theorizing gender perspectives in foreign and second language learning. In R. M. Jiménez Catalán (Ed.), *Gender Perspectives on Vocabulary in Foreign and Second Languages* (pp. 1-19). New York: Palgrave Macmillan.

Terrazas, M., & Llach, M. P. A. (2009). Exploring the increase of receptive vocabulary knowledge in the Foreign Language: A longitudinal study. *International Journal of English Studies, 9*(1), 113-133.

Vidal, K. (2011). A comparison of the effects of reading and listening on incidental vocabulary acquisition. *Language Learning, 61*(1), 219-258.

Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics, 28*(1), 46-65.

Webb, S. (2008). The effects of context on incidental vocabulary learning. *Reading in a Foreign Language, 20*(2), 232-245.

Wilkins, D.A. (1972). *Linguistics in Language Teaching*. London: Arnold.

Xing, P, & Fulcher, G. (2007). Reliability assessment for two versions of Vocabulary Levels Tests. *System, 35*, 182-191.

Appendix 1. Background Questionnaire

CUESTIONARIO

1. Nombre y apellidos:
    _____

2. Sexo: H ☐   M ☐

3. Edad: _____

4. Fecha de nacimiento: _____

5. Lugar de nacimiento: _____
6. Nacionalidad: _____
7. Profesión: _____
8. Curso: _____

9. ¿Cuál es tu lengua materna?, ¿Cuál fue la primera lengua en la que empezaste a comunicarte cuando eras niño: euskera, castellano, ambas o alguna otra?
    _____

10. Lengua dominante de la madre: _____

11. Lengua dominante del padre: _____

12. Lengua (s) que hablabas en casa cuando eras niño: _____
    Con la madre: _____
    Con el padre: _____
    Con los hermanos/hermanas: _____
    Con otros miembros de la familia: _____

13. Lengua (s) que hablaste durante los primeros cinco años de tu vida:

    _____

14.   Lengua (s) extranjeras que has estudiado (incluye el inglés):

	Educación Infantil (cuántas horas a la semana)	**Educación Primaria** (cuántas horas a la semana)	**Educación Secundaria** (cuántas horas a la semana)	**Otras instituciones** (E.O.I., Academia…) (desde cuándo, hasta cuándo y cuántas horas a la semana)
1ª lengua extranjera:	1º: 2º: 3º:	1º: 2º: 3º: 4º: 5º: 6º:	1º: 2º: 3º: 4º:	
2ª lengua extranjera:	1º: 2º: 3º:	1º: 2º: 3º: 4º: 5º: 6º:	1º: 2º: 3º: 4º:	
3ª lengua extranjera:	1º: 2º: 3º:	1º: 2º: 3º: 4º: 5º: 6º:	1º: 2º: 3º: 4º:	
4ª lengua extranjera:	1º: 2º: 3º:	1º: 2º: 3º: 4º: 5º: 6º:	1º: 2º: 3º: 4º:	
Alguna otra:	1º: 2º: 3º:	1º: 2º: 3º: 4º: 5º: 6º:	1º: 2º: 3º: 4º:	

15.   Lenguas que utilizas

    • en casa: _____

    • en el centro educativo al que asistes: _____

    • en el trabajo: _____

    • con los amigos: _____

    • cuando sueñas: _____

16.   Otras lenguas que

    • lees: _____

    • hablas: _____

    • escribes: _____

17.   ¿En qué lenguas te sientes más cómodo en la actualidad?

    _____

18.   ¿Has estado alguna vez en un país de habla inglesa?   Sí ☐   No ☐
    Si tu respuesta es afirmativa, especifica:
    Cuándo: _____
    Dónde: _____
    Duración: _____
    ¿Participaste en algún programa de intercambio durante tu estancia? _____
    ¿Asististe a algún curso de inglés durante tu estancia? _____

19. ¿Has participado alguna vez en alguna colonia donde se hable inglés? Sí ☐ No ☐
    Si tu respuesta es afirmativa, especifica:
    Cuándo: _____
    Dónde: _____
    Duración: _____

¡GRACIAS!

Oxford University Press
and
University of Cambridge Local Examinations Syndicate

Name: …………………………………………………………………………

Date: …………………………………………………………………………

quick placement test

Version 1

This test is divided into two parts:

Part One (Questions 1 – 40) – All students.

Part Two (Questions 41 – 60) – Do not start this part unless told to do so by your test supervisor.

Time: 30 minutes

Part 1

Questions 1 – 5

- Where can you see these notices?
- For questions **1** to **5**, mark **one** letter **A**, **B** or **C** on your Answer Sheet.

1 **Please leave your room key at Reception.**

    A   in a shop
    B   in a hotel
    C   in a taxi

2 **Foreign money changed here**

    A   in a library
    B   in a bank
    C   in a police station

3 **AFTERNOON SHOW BEGINS AT 2PM**

    A   outside a theatre
    B   outside a supermarket
    C   outside a restaurant

4 **CLOSED FOR HOLIDAYS**

Lessons start again on the 8 th January

    A   at a travel agent's
    B   at a music school
    C   at a restaurant

5 **Price per night:**

£10 a tent
£5 a person

    A   at a cinema
    B   in a hotel
    C   on a camp-site

Questions 6 – 10

- In this section you must choose the word which best fits each space in the text below.
- For questions **6** to **10**, mark **one** letter **A**, **B** or **C** on your Answer Sheet.

Scotland

Scotland is the north part of the island of Great Britain. The Atlantic Ocean is on the west and the

North Sea on the east. Some people **(6)** .... Scotland speak a different language called Gaelic.

There are **(7)** .... five million people in Scotland, and Edinburgh is **(8)** .... most

famous city.

Scotland has many mountains; the highest one is called 'Ben Nevis'. In the south of Scotland, there are

a lot of sheep. A long time ago, there **(9)** .... many forests, but now there are only a

(10) ....

Scotland is only a small country, but it is quite beautiful.

6	**A**	on	**B**	in	**C**	at
7	**A**	about	**B**	between	**C**	among
8	**A**	his	**B**	your	**C**	its
9	**A**	is	**B**	were	**C**	was
10	**A**	few	**B**	little	**C**	lot

Questions 11 – 20

- In this section you must choose the word which best fits each space in the texts.
- For questions **11** to **20**, mark **one** letter **A**, **B**, **C** or **D** on your Answer Sheet.

Alice Guy Blaché

Alice Guy Blaché was the first female film director. She first became involved in cinema whilst

working for the Gaumont Film Company in the late 1890s. This was a period of great change in

the cinema and Alice was the first to use many new inventions, **(11)** .... sound and colour.

In 1907 Alice **(12)** .... to New York where she started her own film company. She was

(13) .... successful, but, when Hollywood became the centre of the film world, the best

days of the independent New York film companies were **(14)** .... When Alice died in

1968, hardly anybody **(15)** .... her name.

---

11 **A** bringing   **B** including   **C** containing   **D** supporting

12 **A** moved   **B** ran   **C** entered   **D** transported

13 **A** next   **B** once   **C** immediately   **D** recently

14 **A** after   **B** down   **C** behind   **D** over

15 **A** remembered   **B** realised   **C** reminded   **D** repeated

UFOs – do they exist?

UFO is short for 'unidentified flying object'. UFOs are popularly known as flying saucers,

(16) .... that is often the **(17)** .... they are reported to be. The **(18)** ....

"flying saucers" were seen in 1947 by an American pilot, but experts who studied his claim

decided it had been a trick of the light.

Even people experienced at watching the sky, **(19)** .... as pilots, report seeing UFOs. In

1978 a pilot reported a collection of UFOs off the coast of New Zealand. A television

(20) .... went up with the pilot and filmed the UFOs. Scientists studying this

phenomenon later discovered that in this case they were simply lights on boats out fishing.

16   **A**   because   **B**   therefore   **C**   although   **D**   so

17   **A**   look   **B**   shape   **C**   size   **D**   type

18   **A**   last   **B**   next   **C**   first   **D**   oldest

19   **A**   like   **B**   that   **C**   so   **D**   such

20   **A**   cameraman   **B**   director   **C**   actor   **D**   announcer

Questions 21 – 40

- In this section you must choose the word or phrase which best completes each sentence.
- For questions **21** to **40**, mark **one** letter **A**, **B**, **C** or **D** on your Answer Sheet.

21   The teacher encouraged her students .... to an English pen-friend.

    A   should write   **B**   write   **C**   wrote   **D**   to write

22   They spent a lot of time .... at the pictures in the museum.

    A   looking   **B**   for looking   **C**   to look   **D**   to looking

23   Shirley enjoys science lessons, but all her experiments seem to .... wrong.

    A   turn   **B**   come   **C**   end   **D**   go

24   .... from Michael, all the group arrived on time.

    A   Except   **B**   Other   **C**   Besides   **D**   Apart

25   She .... her neighbour's children for the broken window.

    A   accused   **B**   complained   **C**   blamed   **D**   denied

26   As I had missed the history lesson, my friend went .... the homework with me.

    A   by   **B**   after   **C**   over   **D**   on

27   Whether she's a good actress or not is a .... of opinion.

    A   matter   **B**   subject   **C**   point   **D**   case

28   The decorated roof of the ancient palace was .... up by four thin columns.

    A   built   **B**   carried   **C**   held   **D**   supported

29   Would it .... you if we came on Thursday?

    A   agree   **B**   suit   **C**   like   **D**   fit

30   This form .... be handed in until the end of the week.

    A   doesn't need   **B**   doesn't have   **C**   needn't   **D**   hasn't got

31   If you make a mistake when you are writing, just .... it out with your pen.

A cross   B clear   C do   D wipe

32   Although our opinions on many things .... , we're good friends.

A differ   B oppose   C disagree   D divide

33   This product must be eaten .... two days of purchase.

A by   B before   C within   D under

34   The newspaper report contained .... important information.

A many   B another   C an   D a lot of

35   Have you considered .... to London?

A move   B to move   C to be moving   D moving

36   It can be a good idea for people who lead an active life to increase their .... of vitamins.

A upturn   B input   C upkeep   D intake

37   I thought there was a .... of jealousy in his reaction to my good fortune.

A piece   B part   C shadow   D touch

38   Why didn't you .... that you were feeling ill?

A advise   B mention   C remark   D tell

39   James was not sure exactly where his best interests ....

A stood   B rested   C lay   D centred

40   He's still getting .... the shock of losing his job.

A across   B by   C over   D through

Appendix 3. 1,000 VLT

INSTITUTO: _____

CURSO: _____ **FECHA:** _____

APELLIDOS: _____ **NOMBRE:** _____

Este es un test de vocabulario. En la parte izquierda te presentamos grupos de seis palabras inglesas y a su derecha, la traducción en castellano de sólo tres de ellas. **Escribe** junto a cada traducción, el **número** de la palabra inglesa que tenga el mismo significado. Observa el siguiente ejemplo:

EJEMPLO		RESPUESTA CORRECTA	
1 dog		1 dog	
2 house	_____ negro	2 house	__5__ negro
3 girl	_____ nariz	3 girl	__6__ nariz
4 fork	_____ casa	4 fork	__2__ casa
5 black		5 black	
6 nose		6 nose	

➔

1 could		1 kill	
2 during	_____ podía/pude	2 reply	_____ avanzar
3 this	_____ durante	3 advance	_____ responder
4 piece	_____ para	4 appoint	_____ matar
5 of		5 divide	
6 in order to		6 receive	
1 indeed		1 moment	
2 what	_____ mi	2 separate	_____ separado
3 along	_____ en efecto	3 worse	_____ momento
4 my	_____ algo	4 free	_____ amarillo
5 some		5 heavy	
6 away		6 yellow	
1 church		1 spring	
2 scene	_____ coche	2 danger	_____ hermana
3 hour	_____ problema	3 stone	_____ peligro
4 trouble	_____ hecho	4 product	_____ piedra
5 fact		5 sister	
6 car		6 subject	
1 meet		1 example	
2 leave	_____ poner	2 breadth	_____ anchura
3 put	_____ dar	3 fear	_____ miedo
4 give	_____ utilizar	4 desert	_____ ayuntamiento
5 use		5 bit	
6 begin		6 town hall	
1 wind		1 surround	
2 room	_____ hombre	2 shoot	_____ quedar bien
3 line	_____ línea	3 paint	_____ advertir
4 enemy	_____ noche	4 fit	_____ disparar
5 night		5 command	
6 man		6 warn	

Appendix 4. 2,000 VLT

INSTITUTO: _____

CURSO_____ **FECHA:** _____

APELLIDOS: _____ **NOMBRE:** _____

Este es un test de vocabulario. En la parte izquierda te presentamos grupos de seis palabras inglesas y a su derecha, los significados de sólo tres de ellas. **Escribe** junto a éstos, el **número** de la palabra inglesa correspondiente a dichos significados. Observa el siguiente ejemplo:

EJEMPLO		RESPUESTA CORRECTA	
1 business		1 business	
2 clock	____ part of a house	2 clock	__6__ part of a house
3 horse	____ animal with 4 legs	3 horse	__3__ animal with 4 legs
4 pencil	____ something used for writing	4 pencil	__4__ something used for writing
5 shoe		5 shoe	
6 wall		6 wall	

→

1 coffee
2 disease   _____ money for work
3 justice   _____ a piece of clothing
4 skirt   _____ using the law in the right way
5 stage
6 wage

1 adopt
2 climb   _____ go up
3 examine   _____ look at closely
4 pour   _____ be on every side
5 satisfy
6 surround

1 choice
2 crop   _____ heat
3 flesh   _____ meat
4 salary   _____ money paid regularly for doing a job
5 secret
6 temperature

1 bake
2 connect   _____ join together
3 inquire   _____ walk without purpose
4 limit   _____ keep within a certain size
5 recognize
6 wander

1 cap
2 education   _____ teaching and learning
3 journey   _____ numbers to measure with
4 parent   _____ going to a far place
5 scale
6 trick

1 burst
2 concern   _____ break open
3 deliver   _____ make better
4 fold   _____ take something to someone
5 improve
6 urge

1 attack
2 charm   _____ gold and silver
3 lack   _____ pleasing quality
4 pen   _____ not having something
5 shadow
6 treasure

1 original
2 private   _____ first
3 royal   _____ not public
4 slow   _____ all added together
5 sorry
6 total

1 cream
2 factory   _____ part of milk
3 nail   _____ a lot of money
4 pupil   _____ person who is studying
5 sacrifice
6 wealth

1 ancient
2 curious   _____ not easy
3 difficult   _____ very old
4 entire   _____ related to God
5 holy
6 social