The receptive vocabulary of Spanish 6th-grade primary-school students in CLIL instruction: A preliminary study[1]

El vocabulario receptivo de estudiantes españoles de 6º de educación primaria en contextos AICLE: Un estudio preliminar

Andrés Canga Alonso
University of La Rioja
(Logroño, Spain)

Abstract

One of the key factors in early stages of L2 learning is the number of words learners know. However, there has been little research regarding the receptive vocabulary size of learners involved in content and language integrated learning (CLIL) programmes at primary level. Accordingly, the 2,000-word frequency band of the Vocabulary Levels Test (VLT) (Schmitt, Schmitt, & Clapham, 2001) was administered to explore the receptive vocabulary knowledge of 6th-grade, primary-level Spanish students learning English through CLIL instruction to relate their receptive vocabulary size to their ability to understand written and spoken discourse in English and to establish sex-based differences amongst the participants. The results show that students' receptive vocabulary knowledge is below the 1,000 frequency band, which implies that students may find it difficult to understand spoken and written discourse in English (Laufer, 1992; Nation, 2001; Adolphs & Schmitt, 2004). As for sex-based differences, female students outscored their male partners in the VLT, but these differences were not statistically significant. Nevertheless, further research needs to be conducted with a larger sample of CLIL learners to compare groups from different schools implementing these programmes in the same area, as well as to compare their results with those obtained by students in traditional non-CLIL environments.
Key Words: CLIL; VLT; vocabulary; gender differences; primary education.

Resumen

Uno de los factores clave en el aprendizaje temprano de una L2 es el número de palabras que saben los estudiantes. Sin embargo, hay una carencia de investigación en lo referente al tamaño del vocabulario receptivo de los estudiantes que participan en programas de Aprendizaje Integrado de Lenguas y Contenidos (AICLE) en educación primaria. A la luz de esta falta de investigación se administró el Vocabulary Levels Test (VLT) (Schmitt, Schmitt, & Clapham, 2001) correspondiente a las 2,000 palabras más frecuentes del inglés para explorar el vocabulario receptivo de estudiantes españoles de 6º de primaria que siguen un enfoque AICLE para relacionar su tamaño de vocabulario receptivo con la capacidad de comprensión del discurso hablado y escrito en inglés, y para establecer al mismo tiempo las diferencias de acuerdo con el sexo de los participantes. Los resultados muestran que el tamaño del vocabulario receptivo de los estudiantes está por debajo de la banda de las 1,000 palabras lo que implica que pueden tener dificultades para comprender el

discurso escrito y oral en inglés (Laufer, 1992; Nation, 2001; Adolphs & Schmitt, 2004). En lo que respecta a las diferencias en cuanto al sexo, las chicas obtienen mejores resultados en el VLT que los chicos, pero las diferencias entre los grupos no son significativas estadísticamente. Sin embargo, se necesita llevar a cabo más estudios con una muestra más amplia de alumnos AICLE para poder analizar grupos de diferentes centros educativos de la misma zona y comparar sus resultados con los obtenidos por alumnos en programas no-AICLE.
Palabras Claves: AICLE; VLT; vocabulario; diferencias de género; educación primaria.

INTRODUCTION

In recent decades, the acronym CLIL (Content and Language Integrated Learning) has been used as a generic term to describe all types of approaches in which a second language is used to teach certain content subjects in the curriculum other than language lessons (Ruiz de Zarobe, 2011). The essence of CLIL is integration with a dual focus: "language learning is included in content classes (e.g. maths, history, geography [...], etc.), and content from subjects is used in language learning classes" (Mehisto, Marsh, & Frigols, 2008: p. 11). CLIL also provides real and relevant input for the learner. This input refers to the content that the teacher is presenting as well as the language for classroom management necessary to ensure that learning takes place (Muñoz, 2007).

In line with these assumptions, this paper aims at analysing the English receptive vocabulary size of Spanish 6th Primary school learners of English as a foreign language (EFL) enrolled in a programme in which the content of a curricular subject is taught through English to relate students' receptive vocabulary size to their ability to understand written and spoken discourse in English and to establish sex-based differences among the participants.

CLIL and vocabulary learning

As mentioned above, vocabulary knowledge is acknowledged to be of paramount importance to facilitate students' interaction in the foreign language. Researchers have tackled the issue concerning the number of words necessary to understand spoken discourse (Nation, 2001; Adolphs & Schmitt, 2004) and to read and comprehend texts in the native and foreign language (Anderson & Freebody, 1981; Laufer, 1997). Among the former researchers, Adolphs & Schmitt (2004) estimate that (at least) 2,000 word forms have to be mastered in order to understand around 90% and 94% of spoken discourse in different contexts. Among the latter group, Laufer (1992, 1997) states that a text coverage of 95% can

be reached with a 5,000-word English vocabulary or 3,000-word families, which agrees with the assertions made by Hazenberg & Hulstijn (1996), Nation (1993, 2001) and Cobb & Horst (2004). More recently, Nation (2006) has asserted that 8,000- to 9,000-word families are needed for understanding a written text and a vocabulary of 6,000- to 7,000-word families for comprehension of spoken text, if 98% coverage of a text is desired. Hirsh & Nation (1992) also point out that knowledge of 5,000-word families is necessary to enjoy reading. As we have seen, estimates based on word frequency criteria have been calculated and research claims that gaining command of the 2,000-3,000 most frequent words as soon as possible is vital for the language learner to communicate orally and in written form in the foreign language (Nation, 1993; Nation & Waring, 1997). The sooner the most frequent words are learned by students, the better their language performance will be. As Schmitt (2000: p. 137) claims: "The learning of these basic words cannot be left to chance, but should be taught as quickly as possible, because they open […] the door of further learning".

Assuming that in CLIL settings it is necessary to progress systematically in pupils' content and language learning and using, vocabulary knowledge is of paramount importance in order to favour communication in the classroom. As a result, classroom communication—interaction between peers and teachers—is at the core. There is also growing recognition that dialogic forms of pedagogy—that is, "where learners are encouraged to articulate their learning" (Coyle, Hood, & Marsh, 2010, p. 35)—are potent tools for securing students' engagement and understanding. Focusing teaching on quality discourse understanding between students, and between learners and teachers—where pupils have different opportunities to discuss their own learning with other peers as it progresses, where feedback is integrated into classroom discourse and where apprentices are encouraged to ask as well as answer questions. The challenge in the CLIL setting is that trainees will need to engage in dialogic interactions by using the vehicular language. In this vein, Dalton-Puffer (2007, 2008) states that there are some areas where clear gains are observed in CLIL classrooms (for example: receptive skills, vocabulary, morphology, creativity). These assertion concords with Cummins' (1979, 2008) distinction between basic interpersonal communicative skills (BICS) and cognitive academic language proficiency (CALP), which should be taken into account in order to draw educators' attention to the timelines and challenges that second language learners encounter as they attempt to catch up to their peers in

academic aspects of the school language vocabulary. BICS refers to conversational fluency in a language, while CALP refers to students' ability to understand and express, in both oral and written modes, concepts and ideas that are relevant to success in school. As will be illustrated in the procedure and data gathering section, the 2,000 frequency band of Vocabulary Levels Test (VLT) is based on the frequency lists collected by West (1953), Thorndike & Lorge (1944) list, and Kucera & Francis (1967), so it measures students' ability to recognise academic words from these lists which relates to Cummins' concept of CALP since learners should be able to recognise and understand the meaning of the terms provided and relate them to their definitions. Consequently, a link between CLIL instruction and vocabulary learning can be established.

In recent decades, a considerable number of studies have investigated receptive vocabulary size or the number of words a learner knows. Most studies coincide in indicating that vocabulary size grows as proficiency level in the foreign language (Barrow, Nakanishi, & Ishino, 1999; Fan, 2000), exposure to the target language (Schmitt, 1998; Golberg, Paradis, & Crago, 2008), or frequency of input (Vermeer, 2001) increase. Moreover, this gain follows a systematic order related to frequency, since at the lowest levels of proficiency learners are familiar with the most frequent words, but as their experience with the foreign language increases, less frequent words are incorporated into the lexicon (Barrow *et al.*, 1999; Vermeer, 2001; Milton, 2009). The probability of a word being known by foreign language learners rises with its frequency, so higher-frequency words have a greater possibility of being known. It seems evident that a content-based approach provides more opportunities to learn, either explicitly or implicitly, target vocabulary in meaningful situations (Muñoz, 2007; Pérez-Vidal, 2009), since learners are exposed to the target language for a longer period than are students' enrolled in traditional EFL classrooms. Several studies have explored this relationship between time of exposure and vocabulary learning. Thus, Xanthou (2010) states that CLIL has a positive impact in a group of primary school children in Cyprus regarding students' vocabulary tests results, which demonstrates that by attaching words to their surroundings, the likelihood of comprehension and retention is increased. These gains in the size of receptive vocabulary are in line with other research conducted in Spain (Jiménez Catalán, Ruiz de Zarobe, & Cenoz, 2006; Jiménez Catalán & Ruiz de Zarobe, 2009), where significant results were obtained in favour of the CLIL group in receptive

vocabulary knowledge. Nevertheless, these studies have not focused on sex-based differences inside the CLIL group: therefore, the present study aims at shedding some light on sex-based differences regarding students' receptive vocabulary knowledge at the end of primary education in a school in the north of Spain.

Table 1 presents a summary of previous estimates of receptive vocabulary size of L2 learners of English at primary and secondary level in CLIL and non-CLIL classrooms in Spain. Although the aim of the present study is to analyse 6th-grade CLIL primary students' receptive vocabulary size, due to the scarcity of research in this field, it was considered relevant to include research conducted in non-CLIL secondary-school settings in the same educational context to compare both groups of students. Studies have been reported in which learners received a number of hours of instruction similar to that of the students analysed in the present research. These studies are ordered according to the receptive vocabulary size of learners. As can be seen, the results obtained show considerable differences in receptive vocabulary knowledge on the part of the learners who were investigated. L2 students' vocabulary knowledge figures are also complex to compare due to differences concerning pupils, and their contexts of learning. However, it is useful to compare the results presented in Table 1 with those obtained by students in the current study since, insofar as it has been possible to determine, there are not many studies in Spain that have related CLIL with receptive vocabulary size at primary school level apart from that of Jiménez Catalán and Ruiz de Zarobe (2009).

Table 1. Average receptive vocabulary size.

Study	Receptive Vocabulary Size	Hours of Instruction	Participants' learning context
Agustín Llach & Terrazas Gallego (2012)	1206	944	secondary education (3rd ESO/9th grade non-CLIL)
Terrazas Gallego & Agustín Llach (2009)	817 words	734	secondary education (1st ESO/7th grade non-CLIL)
Jiménez Catalán & Ruiz de Zarobe (2009)	800 words	960	primary education (6th grade CLIL)
Jiménez Catalán & Terrazas Gallego (2008)	559 words	419	primary education (4th grade non-CLIL)
Agustín Llach & Terrazas Gallego (2012)	663 words	629	primary education (6th grade, non-CLIL)

The results in Table 1 also indicate that non-CLIL primary students' receptive vocabulary knowledge is clearly below 1,000 words and ninth-graders in non-CLIL instructional settings with a similar amount of instruction in the foreign language as the students in the present study's sample have a receptive vocabulary size greater than 1,000 words.

Having analysed the importance of learning vocabulary in L2 and CLIL instruction, as well as having referred to the research conducted on receptive vocabulary knowledge in non-CLIL contexts in Spain, the importance of sex-based differences in the literature on vocabulary acquisition can be considered. The role of sex has occupied an outstanding place in current research on vocabulary acquisition. Receptive and productive vocabulary knowledge of male and female learners has been widely examined, and scholars have reached different conclusions. Boyle (1987) concludes that, exceptionally, boys are superior to girls in the comprehension of heard vocabulary. Similarly, Scarcella and Zimmerman (1998) find that men perform significantly better than women in a test of academic vocabulary recognition, understanding, and use. In Lynn, Fergusson, and Horwood (2005) and in Edelenbos and Vinjé (2000), males also outperformed females with respect to vocabulary knowledge in the foreign language. By contrast, in Nyikos' study (1990), women performed better than men in a memorisation test of German vocabulary. Nevertheless, Jiménez Catalán and Terrazas Gallego (2008) discover no significant sex-based differences in performance on a receptive vocabulary test implemented with primary students. In a recent longitudinal study on vocabulary knowledge and gender differences at primary and secondary level, Agustín Llach and Terrazas Gallego (2012) obtain similar results, finding very slight differences concerning receptive vocabulary knowledge among males and females across grades in the context of Spanish primary education. These two latter studies used the same test as that implemented in the present study, the Vocabulary Levels Test (VLT) (Schmitt, Schmitt, & Clapham, 2001), to measure students' receptive vocabulary knowledge. These same studies are, therefore, referred to in the following sections.

A set of recent studies compiled in Jiménez Catalán (2010) also point to mixed results regarding gender differences and the acquisition, development, meaning, and use of vocabulary by adult, adolescents, and young learners of English and Spanish in Spain, Canada, and the USA. Jiménez Catalán moreover

relates learner gender and meaning, vocabulary use, lexical creation, lexical production, and word association, as well as how words encode both patterns of gender representation and gender identities.

Finally, in a lexical availability tests study in which students had to respond to 15 cues, highly significant differences are found in favour of females in the mean number of words produced (Jiménez Catalán & Ojeda Alba, 2009). Therefore it can be concluded that the relationships between vocabulary and gender are not enduring but may be specific to context and test type, as well as perhaps influenced by L1, age, or L2 proficiency (Sunderland, 2010).

Considering the aforementioned studies, it can be said that there is a scarcity of research that correlates CLIL instruction with sex-based differences amongst Spanish primary-school students—such as is essayed in the present study. For these reasons, this study aims at investigating the receptive vocabulary knowledge of sixth-grade Spanish primary school students learning English through a CLIL approach to relate their receptive vocabulary size to their ability to understand written and spoken discourse in English (Nation, 1993, Laufer, 1992; Hazenberg & Hulstijn, 1996; Laufer, 1997; Nation, 2001; Adolphs & Schmitt, 2004; Cobb & Horst, 2004).

METHODOLOGY

Participants

The sample for the study consisted of 79 students. The group was made up of 46 boys and 33 girls enrolled in a CLIL programme in a primary school in the north of Spain. All the participants were in the sixth grade of primary education. The sample was homogeneous with regards to social environment since all the students lived in the same area. Students also shared Spanish as their mother tongue (L1), and they were 11-12 years old. The participants received instruction in English in two curricular subjects: English as a foreign language (EFL), and Natural Sciences. They were exposed to the foreign language for a total amount of 944 hours.

Procedures and data collection

The 2,000-word frequency band (2k) from the receptive version of the Vocabulary Levels Test (VLT) was used to measure the receptive vocabulary size of these

subjects (Schmitt, Schmitt, & Clapham, 2001, version 2). This test is based on the frequency lists collected by West (1953) in the General Service List and the Thorndike & Lorge (1944) list, which were checked against the list compiled by Kucera & Francis (1967), known as the Brown Corpus.

In the 2k VLT (Schmitt, Schmitt, & Clapham, 2001), test-takers have to match a target word with the corresponding definition. A total of 60 target words are used for testing. Ten groups of 6 words and 3 definitions make up the test. Each correct answer (matching each target word with its definition) is given 1 point, so that the maximum score of the test is 30 points. The research studies that have reported on the validity and reliability of the 2k VLT (Beglar & Hunt, 1999; Read 2000) show that the test is not only valid and consistent in its measurements but also that, in fact, it measures what it sets out to measure.

The author of this paper collected the data in 1 session during class time. The time allotted to complete the task was 10 minutes. At the beginning of the test, clear instructions, together with an example, were given both orally and in written form in the students' mother tongue to clarify what they were being asked to do.

In order to calculate students' word estimates, Nation's formula "Vocabulary size = *N* correct answers multiplied by total *N* words in dictionary (the relevant word list) divided by *N* items in test" (Nation, 1990, p. 78) was applied.

The sample was also analysed by means of SPSS 19 to test whether there were statistically significant differences between the boys and the girls who participated in the study. Thus, the Shapiro-Wilk and Kolmogorov-Smirnov statistical tests were implemented to prove whether the sample met the normality assumption. As will be shown in the results section, the sample did not meet this assumption, so non-parametric tests had to be implemented. The Mann-Whitney U test was then chosen to perform inferential statistics amongst male and female learners. This test is used to compare differences between two independent groups when the dependent variable is either ordinal or interval/ratio but not normally distributed. In the present study, the Mann-Whitney U test was used to measure whether students' VLT scores differed based on sex.
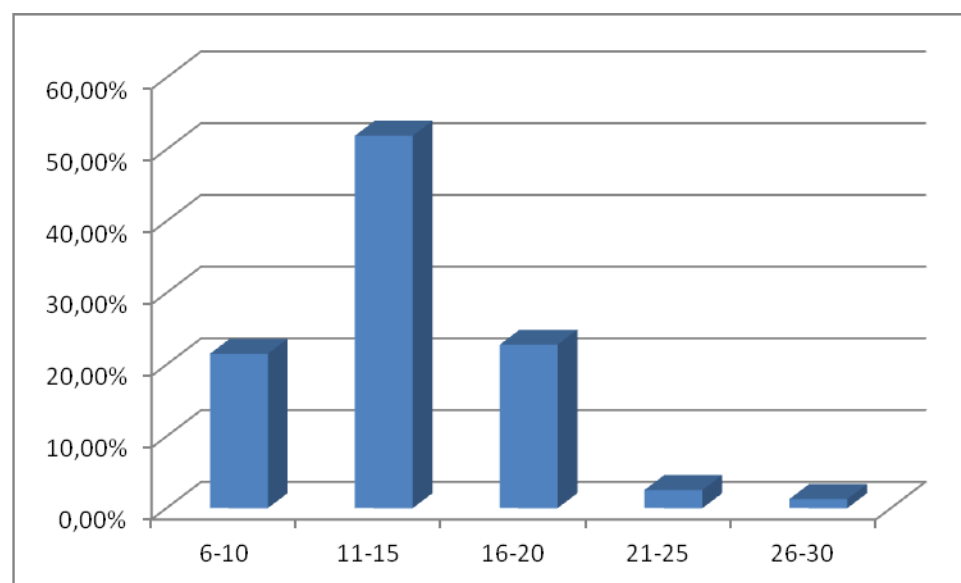
RESULTS

As can be seen in Table 2, the maximum score in the sample was 29 points out of 30, which was attained by one of the students, whereas the minimum score (6 out of 30) was achieved by 3 participants. These figures indicate that students are a long way from having learnt the 2,000 most frequent words, according to the frequency lists collected by Thorndike & Lorge (1944), West (1953), and Kucera & Francis (1967). These data also imply that students might have problems understanding spoken and written discourse in English, since they need to have gained at the least the 2,000 most frequent words in order to communicate orally and in written form in the foreign language (Nation, 1993; Nation & Waring, 1997).

Table 2. VLT 2,000 results.

	Min.	Max.	Means	SD
n=61	6	29	13.54	4.17

This profile is illustrated in the rankings of percentages summarized in Figure 1. The results show that 21.52% of our informants scored between 6 and 10 points, 51.9% scored between 11 and 15, while 22.78% scored between 16 and 20 points, 2.53% between 21 and 25, and 1.27% between 26 and 30.
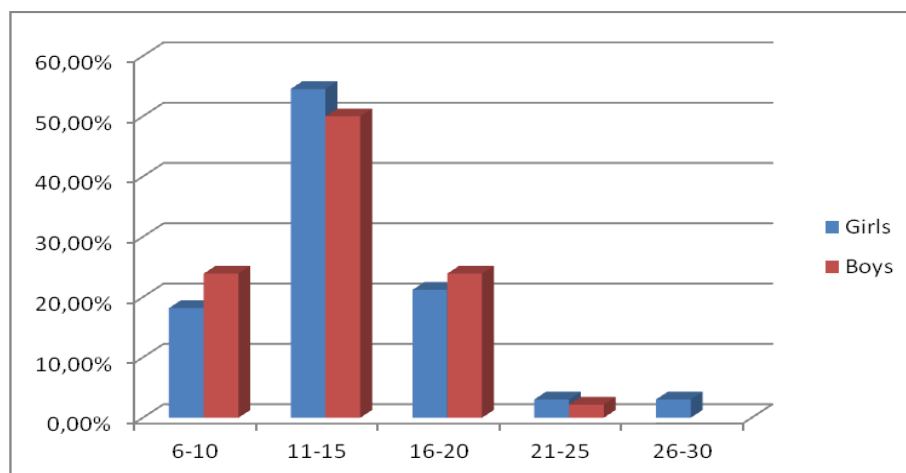
Figure 1. Frequency distribution of tests scores.

Students' scores were translated into the number of known words for each frequency level applying Nation's formula: "Vocabulary size = N correct answers multiplied by total N words in dictionary (the relevant word list) divided by N items in test" (Nation, 1990, p. 78). The mean obtained (903 words) confirms the previous presupposition: that the learners in this sample did not know all of the 1,000 most frequent words in English. However, this result is similar to previous studies conducted with primary school students in Spain and abroad, where 400-800 hours of instruction have been shown to lead to vocabulary sizes of around 1,000 words (Staehr, 2008; Jiménez Catalán & Ruiz de Zarobe, 2009; Terrazas & Agustín Llach, 2009). This hypothesis is reinforced by the fact that 37.98% of the participants have a receptive vocabulary size equal to or higher than 1,000 words in the 2k VLT.
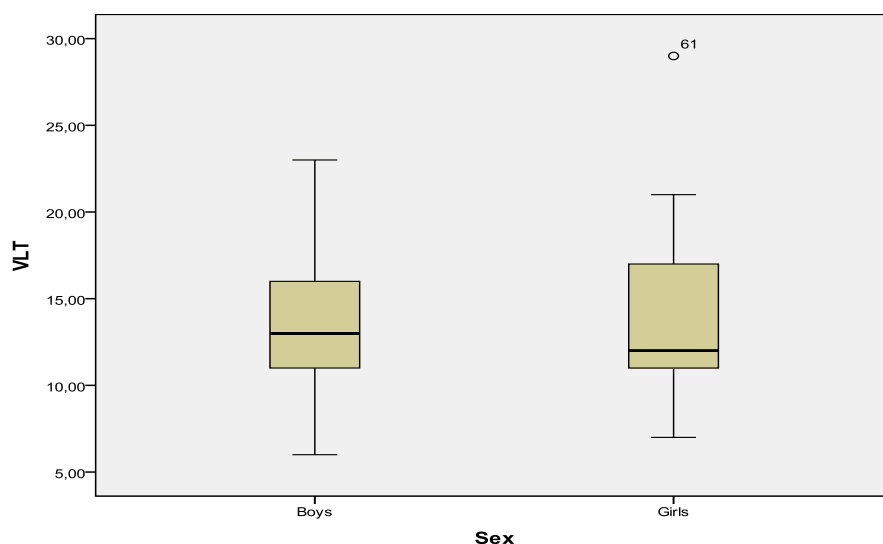
As for differences between the sexes, girls obtained better maximum (29 points *vs.* 23 points) and minimum scores (7 points *vs.* 6 points) in the test. Therefore, the mean (13.88) is also higher for the female participants, which indicates that their overall scores were better. Figure 2 compares the frequency distribution of test scores attained by male and female learners. As can be seen, girls attained higher medium scores, since more than 54% scored between 11-15 points in the VLT, as opposed to 50% of the boys. In contrast, boys obtained better results in the 16-20 frequency band (23.9% vs. 21.1%). Finally, girls also outperformed their male counterparts in the 21-25 and 26-30 frequency bands.

Figure 2. Frequency distribution of tests scores for boys and girls.

The box-plot in Figure 3 illustrates the median values for boys and girls, which concord with the data analysed above.

Figure 3. Median vales according to sex.

Regarding word estimates and sex-based differences, Table 3 illustrates that mean values and maximum scores are higher for girls, whereas the lowest scores were obtained by 3 boys. These findings also imply that word estimate values are also higher for female students.

Table 3. Distribution of word estimates according to sex.

	Min	Max	Mean
Boys (n=46)	400	1533	887
Girls (n=33)	467	1933	925

Kolmogorov-Smirnov and Shapiro-Wilk parametric tests were implemented in order to ascertain whether the sample met the normality assumption. As shown in Table 4, the p-values obtained were lower than p=0.05 for the girls, so the sample did not meet normality and, therefore, non-parametric tests were applied.

Table 4. Parametric tests for sex-based differences.

	Sex	Kolmogorov-Smirnov			Shapiro-Wilk		
		Statistics	gl	Sig.	Statistics	gl	Sig.
VLT	Boys	0.071	46	0.200	0.984	46	0.758
	Girls	0.175	33	0.012	0.909	33	0.009

The U Mann-Whitney test was conducted to calculate inferential statistical differences among the groups. Its results reveal that there are no significant sex differences at a significance level of 5% (p=0.90) in vocabulary size estimations. Table 5 offers these results.

Table 5. Results of inferential statistics for sex-based differences.

n=79	VLT
U Mann-Whitney	71.500
p (two-tailed)	0.90

DISCUSSION

The findings of the present study reveal that students' receptive vocabulary knowledge is lower than 1,000 words. These results are higher than those attained by learners of a similar age who have received similar amounts of instruction in EFL in the same educational context (Jiménez Catalán & Ruiz de Zarobe, 2009; Terrazas Gallego & Agustín Llach, 2009). Thus, after 944 hours of instruction, participants in the present study had mean score for word estimates of 903 words, whereas the word estimates for Jiménez Catalán and Ruiz de Zarobe's (2009) informants was 800 words after being exposed to English for a total amount of 960 hours. The students in the present study also had a higher mean score than that obtained by Terrazas Gallego and Agustín Llach (2009) and Agustín Llach and Terrazas Gallego (2012) with first-grade secondary-school students (817 words) and sixth-grade primary students of the same age (663 words) enrolled in a non-CLIL instructional programme. In contrast, the present study reveals poorer results if compared to a group of third-grade secondary-school students who have been exposed to English for the same number of hours of instruction in a non-CLIL environment in the same context, since the mean score for this group of students is 1206 words (Agustín Llach & Terrazas Gallego, 2012). These data concord with previous research on age and vocabulary acquisition in different contexts, since older students tend to achieve higher vocabulary profiles in comparison with younger students who have received a similar number of hours of instruction in English as a foreign language (Cenoz, 2002; Cummins & Swain, 1986; Lasagabaster & Doiz, 2003; Miralpeix, 2006).

These poorer scores could also be explained by reference to the kind of vocabulary input to which students are exposed in their CLIL classroom and in their textbooks. This vocabulary input favours cognitive academic language proficiency (CALP) since it focuses on aspects related to fauna, flora, cells, ecosystems, population, and history. However, as Bruton (2011) states, there are numerous anomalies not only in the implementation of CLIL programmes but also in the methodology CLIL teachers implement in their daily practice. Therefore, it seems necessary for teachers to revise their methodologies in order to focus on the type of academic vocabulary input that is included in the textbooks they most often use in the classroom, as well as the number of occurrences of the words contained in them, so that their students can progressively acquire new words (Jiménez Catalán & Mancebo Francisco, 2008) related to the topics they are dealing with in their content lessons.

On the other hand, the wordlists used to measure students' receptive vocabulary size (Thorndike & Lorge, 1944; West 1953; Kucera & Francis, 1967) are not adapted to CLIL instruction, which might imply that some of the words learnt by our informants in their content lessons may not be reflected in the most frequent 2,000 English words.

The results also show that the students in the present study may find it difficult to understand written and spoken discourse, since their mean word estimate is lower than 1,000 words, and they would need to master at least 2,000 word forms to be able to understand around 90% to 94% of spoken discourse in different contexts (Nation, 2001; Adolphs & Schmitt, 2004) and have about a 5,000-word English vocabulary or 3,000 word families to reach a text coverage of 95% (Laufer, 1992; Hazenberg & Hulstijn, 1996; Laufer, 1997; Nation, 2001; Adolphs & Schmitt, 2004; Cobb & Horst, 2004).

As for sex-based differences, the data do not reveal statistically significant differences between the boys and girls in the sample. Nevertheless, the results indicate that girls performed slightly better than boys in the test administered. These outcomes agree with the results of research conducted in similar contexts and age groups (Jiménez Catalán & Terrazas Gallego, 2008; Terrazas Gallego & Agustín Llach, 2009; Agustín Llach & Terrazas Gallego, 2012), where girls slightly outperformed boys. Consequently, we can conclude that the relationships between vocabulary and sex are not enduring, though this assumption should be taken with care since the sample of CLIL learners is quite small.

Conclusions and further study

In the light of the present results, while the receptive vocabulary size of the learners in the present study is lower than 1,000 words, the findings also reveal that CLIL learners obtained better results than non-CLIL learners of the same age and educational context. Therefore, CLIL instruction seems to be positive for this group of learners since their gains in receptive vocabulary are higher when compared to students of their same age and academic background.

On the other hand, no statistically significant sex-based differences were found amongst the students in the sample, although the girls attained higher scores than their male partners in the Vocabulary Levels Test. These data also concord with previous findings regarding non-CLIL students, where girls slightly outperformed boys (Jiménez Catalán & Terrazas Gallego, 2008; Terrazas Gallego & Agustín Llach, 2009; Agustín Llach & Terrazas Gallego, 2012). This may indicate that girls' receptive vocabulary size is larger than boys', regardless of the type of instruction (CLIL or non-CLIL).

Nevertheless, these findings should be treated with caution, since the study presents limitations based on the relatively small number of participants. It is worth mentioning that it has been difficult to collect data from larger samples of CLIL students since not many schools in the region where the study was conducted are currently implementing CLIL programmes in comparison with the number of institutions whose syllabi are based on traditional EFL (non-CLIL) instruction.

Further research needs to be carried out with larger samples of CLIL learners to compare groups from different schools implementing these programmes in the same area, as well as comparing their results with those obtained by students in traditional, non-CLIL environments. Longitudinal studies with CLIL learners would also be helpful to determine whether receptive vocabulary growth is incremental throughout primary school instruction and to test whether sex-based differences remain the same or become statistically significant in girls' favour or, contrariwise, whether boys obtain significantly better results than their female classmates as students grow older. Receptive vocabulary size can also be attached to the six levels of the Common European Framework of Reference (2001); therefore X-Lex (Meara and Milton, 2003; Milton,

2010) could be implemented to test whether there are differences in students' level according to type of instruction and sex. In addition, discrete receptive tests based on frequency (such as the 2,000 VLT) should be complemented by specific tests on the vocabulary and terms related to the content subject studied in English by the CLIL students. This would provide a more accurate picture of learners' vocabulary differences in CLIL and non-CLIL situations.

Finally, productive vocabulary should also be assessed by means of composition tasks to ascertain whether CLIL learners also outscore non-CLIL students of the same age and educational background as seems to be the case with receptive vocabulary knowledge.

REFERENCES

Adolphs, S. & Schmitt, N. (2004). Vocabulary coverage according to spoken discourse context. In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a second language* (pp. 39-52). Amsterdam, the Netherlands: John Benjamins.

Agustín Llach, M. P. & Terrazas Gallego, M. (2012). Vocabulary knowledge development and gender differences in a second language. *ELIA*, *12*, 45-75.

Anderson, R. C. & Freebody, P. (1981). Vocabulary knowledge. In J. Guthrie (Ed.), *Comprehension and teaching: Research reviews* (pp. 203-226). Newark, DE: International Reading Association.

Barrow, J.; Nakanishi, Y., & Ishino, H. (1999). Assessing Japanese college students' vocabulary knowledge with a self-checking familiarity survey. *System*, *27*, 223-247.

Beglar, D. & Hunt, A. (1999). Revising and validating the 2000 word level and university word level vocabulary tests. *Language Testing*, *16*, 131-62.

Boyle, J. P. (1987). Sex differences in listening vocabulary. *Language Learning*, *37*(2), 273-284.

Bruton, A. (2011). Is CLIL so beneficial or just selective? Re-evaluating some of the research. *System*, *39*, 523-532.

Cenoz, J. (2002). Age differences in foreign language learning. *ITL Review of Applied Linguistics*, *135-136*, 125-142.

Cobb, T. & Horst, M. (2004). Is there room for an academic word list in French? In P. Bogaards and B. Laufer (Eds.), *Vocabulary in a second language* (pp. 15-38). Amsterdam, the Netherlands: John Benjamins.

Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment.* Cambridge, England: Cambridge University Press.

Coyle, D., Hood, P., & Marsh, D. (2010). *CLIL: Content and language integrated learning.* Cambridge, England: Cambridge University Press.

Cummins, J. (1979). Cognitive/academic language proficiency, linguistic interdependence, the optimum age question and some other matters. *Working Papers on Bilingualism, 19*, 121-129.

Cummins, J. (2008). BICS and CALP: Empirical and theoretical status of the distinction. In B. Street & N. H. Hornberger (Eds.), *Encyclopedia of Language and Education* (2nd ed., Vol. 2, pp. 71–83). New York City, NY: Springer.

Cummins, J. & Swain, M. (1986). Linguistic interdependence: A central principle of bilingual education. In J. Cummins & M. Swain (Eds.), *Bilingualism in Education. Aspects of the theory, research and practice* (pp. 80-95). London, England: Longman.

Dulton-Puffer, Ch. (2007). *Discourse in content and language integrated learning (CLIL).* Amsterdam, the Netherlands/Philadelphia, PA: John Benjamins.

Dulton-Puffer, C. (2008). Outcomes and processes in content and language integrated learning (CLIL): Current research from Europe. In W. Delanoy & L. Volkmann (Eds.), *Future perspectives for English language teaching* (pp. 139-157). Heidelberg, Germany: Carl Winter.

Edelenbos, P. & Vinjé, M. P. (2000) The assessment of a foreign language at the end of primary (elementary) education. *Language Testing, 17*(2), 144-162.

Fan, M. (2000). How big is the gap and how to narrow it? An investigation into the active and passive vocabulary knowledge of L2 learners. *RELC Journal, 31*, 105-119.

Golberg, H., Paradis, J., & Crago, M. (2008). Lexical acquisition over time in minority first language children learning English as a second language. *Applied Psycholinguistics, 29*, 41–65.

Hazenberg, S. & Hulstijn, J. H. (1996). Defining a minimal receptive second-language vocabulary for non-native university students: An empirical investigation. *Applied Linguistics, 17*, 145–63.

Hirsh, D. &. Nation, P. S. N. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language, 8*, 689-696.

Jiménez Catalán, R. M. (ed.). (2010). *Gender perspectives on vocabulary in foreign and second language.* Basingstoke: Palgrave.

Jiménez Catalán, R. M. & Mancebo Francisco, R. (2008). Vocabulary input in EFL textbooks. *RESLA*, *21*, 147-165.

Jiménez Catalán, R. M. & Ojeda Alba, J. (2009). Girls' and boys' lexical availability in EFL. *ITL - International Journal of Applied Linguistics*, *158*, 57-76.

Jiménez Catalán, R. M. & Ruiz de Zarobe, Y. (2009). The receptive vocabulary of EFL learners in two instructional contexts: CLIL versus non-CLIL. In R. Mª Jiménez Catalán & Ruiz de Zarobe, Y. (Eds.), *Content and language integrated learning. Evidence from research in Europe* (pp. 81-92). Bristol, England: Multilingual Matters.

Jiménez Catalán, R. M., Ruiz de Zarobe Y., & Cenoz, J. (2006). Vocabulary profiles in English as a subject and as a vehicular language. *VIEWZ*, *15*(3), 23-27. Retrieved from http://anglistik.univie.ac.at/fileadmin/user_upload/dep_anglist/weitere_Uploads/Views/views15_3_clil_special.pdf

Jiménez Catalán, R. M. & Terrazas Gallego, M. (2008). The receptive vocabulary of English foreign language young learners. *Journal of English Studies*, *5*, 173-191.

Kucera, H. & Francis, W. N. (1967). *A computational analysis of present day American English.* Providence, RI: Brown University Press.

Lasagabaster, D. & Doiz, A. (2003). Maturational constraints on foreign language written production. In M. P. García Mayo & M. L. García Lecumberri (Eds.), *Age and the acquisition of English as a foreign language* (pp. 136-160). Clevedon, England: Multilingual Matters.

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, England: Macmillan.

Laufer, B. (1997). The lexical plight in second language reading: Words you don't know, words you think you know, and words you can't guess. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 20-34). Cambridge, England: Cambridge University Press.

Lynn, R., Fergusson, D. M., & Horwood, L. J. (2005). Sex differences on the WISC-R in New Zealand. *Personality and Individual Differences*, *39*, 103-114.

Mehisto, P., Marsh, D., & Frigols, M. J. (2008) *Uncovering CLIL: Content and language integrated learning in bilingual and multilingual contexts.* Oxford, England: Macmillan.

Meara, P. & Milton, J. (2003). *X-Lex: The Swansea levels test.* Newbury, England: Express Publishing.

Milton, J. (2009). *Measuring second language vocabulary acquisition.* Bristol, England: Multilingual Matters.

Milton, J. (2010). The development of vocabulary breadth across the CEFR Levels: A common basis for the elaboration of language syllabuses, curriculum guidelines, examinations, and textbooks across Europe. In I. Bartning, M. Martin and I. Vedder (Eds.), *Communicative proficiency and linguistic development: Intersections between SLA and language testing research.* Eurosla Monograph, 1, 211-232. Retrieved from http://eurosla.org/monographs/EM01/211-232Milton.pdf

Miralpeix, I. (2006) Age and vocabulary acquisition in English as a foreign language (EFL). In C. Muñoz (ed.), *Age and rate of foreign language learning* (pp. 89-106). Clevedon, England: Multilingual Matters.

Muñoz, C. (2007). CLIL: Some thoughts on its psychological principles. *RESLA, Vol. Monográfico,* 11-17.

Nation, P. S. N. (1990). *Teaching and learning vocabulary.* New York City, NY: Newbury.

Nation, P. S. N. (1993). Using dictionaries to estimate vocabulary size: Essential, but rarely followed, procedures. *Language Testing, 10,* 27-40.

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

Nation, P. S. N. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review/La revue canadienne des langues vivantes, 63,* 59-81.

Nation, P. S. N. & 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, England: Cambridge University Press.

Nyikos, M. (1990). Sex-related differences in adult language learning: Socialization and memory factors. *Modern Language Journal, 74,* 273-287.

Pérez-Vidal, C. (2009). The integration of content and language in the classroom: A European approach to education (the second time around). In E. Dafouz & M. C. Guerini, (Eds.), *CLIL across educational levels* (pp. 3-16). Madrid, Spain: Richmond.

Read, J. (2000). *Assessing vocabulary*. Cambridge, England: Cambridge University Press.

Ruiz de Zarobe, Y. (2011). Which language competencies benefit from CLIL? An insight into applied linguistics research. In Y. Ruiz de Zarobe, J. M. Sierra & F. Gallardo del Puerto (Eds.), *Content and foreign language integrated learning* (pp. 123-153). Bern, Switzerland: Peter Lang.

Scarcella, R. & Zimmerman, C. (1998). Academic words and gender. ESL student performance on a test of academic lexicon. *Studies in Second Language Acquisition*, *20*, 27-49.

Schmitt, N. (1998). Tracking the incremental acquisition of second language vocabulary: A longitudinal study. *Language Learning*, *48*, 281-317.

Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge, England: Cambridge University Press.

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

Staehr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing. *Language Learning Journal*, *36*, 139-52.

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. 1-22). Basingstoke, England: Palgrave.

Terrazas Gallego, M. & Agustín Llach, M. P. (2009). Exploring the increase of receptive vocabulary knowledge in the foreign language: A longitudinal study. *International Journal of English Studies*, *9*(1), 113-133. Retrieved from http://revistas.um.es/ijes/ article/viewFile/90681/87481

Thorndike, E. & Lorge, I. (1944). *The teacher's word book of 30,000 words*. New York City, NY: Teachers College, Columbia University.

Vermeer, A. (2001). Breadth and depth of vocabulary in relation to L1/L2 acquisition and frequency of input. *Applied Psycholinguistics*, *22*, 217-34.

Xanthou, M. (2010). Current trends in L2 vocabulary learning and instruction. Is CLIL the right approach? In A. Psaltou-Joycey & M Mattheoudakis (Eds.), *14th International Conference: Advances in Research on Language Acquisition and Teaching – Selected Papers* (pp. 459-471). Thessaloniki, Greece: Greek Applied Linguistics Association (GALA). Retrieved from http://www.enl.auth.gr/gala/14th/Papers/English%20papers/Xanthou.pdf

West, M. (1953). *A general service list of English words.* London, England: Longman.

BIODATA

Andrés CANGA ALONSO is a lecturer at the University of La Rioja (Spain). He was also a secondary-school EFL teacher for 8 years. His research focuses on applied linguistics, especially on vocabulary acquisition in EFL and CLIL instruction. He is also interested in the development of learner autonomy by means of the ELP and competence-based approaches. He is member of GLAUR research group and Centre of Research in Applied Linguistics (CRAL) at the University of La Rioja.