



Receptive vocabulary measures for EFL Costa Rican high school students

DAMARIS CASTRO-GARCÍA*
Universidad Nacional (Costa Rica)

Received: 05/08/2016. Accepted: 13/02/2017.

ABSTRACT

The study offers a glimpse of the current situation of foreign language education in the Costa Rican context from the perspective of vocabulary knowledge, particularly passive vocabulary size. Students from two institutions participated: one school implements Content Based Teaching while the other follows traditional, Foreign Language Teaching instruction. This research aims to describe the receptive vocabulary profile of students and to compare the vocabulary levels of students between two gender groups and under two types of English language teaching. These measures are established following the idea originally presented in Paul Nation's (1983, 1990) Vocabulary Levels Test. In this case, Schmitt, Schmitt and Clapham's (2001) Version 2 test was used to define passive vocabulary levels. Finally, the results of this analysis are compared to results for similar population samples in other studies.

KEYWORDS: receptive vocabulary, secondary students, vocabulary size, Vocabulary Levels Test, Costa Rica.

1. INTRODUCTION

Although native tongue vocabulary appears to develop effortlessly and quickly in beginning stages, second language vocabulary learning requires much more careful attention. Research shows that students' vocabulary size is linked directly to their ability to complete different tasks: from basic oral communication to reading novels in the target language. These and several other features of vocabulary lend paramount importance to vocabulary studies. Given that vocabulary knowledge serves as a cornerstone for L2 acquisition and that it has a direct impact on students' learning of the second language, it is one of the key elements for language teachers and researchers tracking students' language progress. Meara (2009) insists

**Address for correspondence:* Damaris Castro-García. Universidad Nacional, Escuela de Literatura y Ciencias del Lenguaje, Heredia, Costa Rica; e-mail: dcastro@una.cr.

on the significance of acknowledging the differences between native and non-native speakers concerning vocabulary acquisition, degree of knowledge and use. How then should teachers and foreign language learners deal with vocabulary acquisition? Schmitt (2000) speaks both of the need for instruction in basic vocabulary and in developing vocabulary-learning strategies to aid students in strengthening this learning process. Pigada and Schmitt (2006) argue in favor of extensive reading as an effective way to acquire vocabulary. They add that reading can enhance vocabulary acquisition in terms of spelling, meaning and grammar. Nation (2006) mentions that whereas the 2,000 high-frequency words should be assigned classroom time to be learned, the situation is different for low-frequency words, for which he recommends that instructors spend time teaching strategies such as guessing words from context, using flash cards, memory or a dictionary for students to learn these words outside the classroom. Nation emphasizes the importance of repetition in the learning process and insists that “learners not only need to gradually meet the most frequent 9,000 word families, but they have to meet them often enough to have a chance of learning them” (2014: 2). Clearly, many different options exist and learners may come up with their own effective ways of learning vocabulary, but why actually is vocabulary so important?

Despite a recent increase in studies dealing with vocabulary gain and use (see Laufer, 1989, 1992; Milton, 2010; Nation, 1983, 1990, 2001, 2006; Read, 1988, 2007; among others), this area of second language acquisition has traditionally received limited attention in research. In recent decades, progress has been made on different fronts lending vocabulary a more prominent role. Referential scholars point to the role of vocabulary as a keystone in second language (L2) learning. For example, for Read (2000: 1), “words are the basic building blocks of language, the units of meaning from which larger structures such as sentences, paragraphs and whole texts are formed”. Aitchison (2012: 53) highlights the importance of vocabulary: “[w]ords...are precision instruments which should be used with care and accuracy”. Nation (2001) ascribes a great deal of importance to high-frequency words, and insists that they play such an essential role in learners’ language that both instructors and learners alike should invest time in them. For Nation (2001: 16), the “frequency, coverage and range” of these words validates dedicating time to them inside and outside of the classroom. Although the term vocabulary has been used above as a broad term encompassing passive and active vocabulary regardless of its depth or breadth, this study will concentrate only on the analysis of receptive vocabulary measures. For that purpose, the notion presented by Heaton (1990: 79) regarding passive or receptive vocabulary as that which “you expect your students to recognize” should be kept in mind for the remainder of this paper.

Concerning vocabulary measures, Nation and Webb (2011: 245) point to how “[m]easures of lexical richness should allow us to distinguish between the language of more and less proficient learners”. Having access to this information provides educators with the tools required to promote practices that can assist learners in the language learning process,

and help them advance in proficiency. Nation (2006: 494) argues that “[t]he most frequent 2,000 words of English, which include most of the function words, are the essential widely used words of a language”. Read (1988), following Nation (1983, 1990), considers that the most frequent words in a language are the first to be learned by students and that any knowledge of vocabulary pertaining to upper vocabulary levels will mean, by default, that students know the preceding vocabulary levels. According to this author (see also Schmitt, Schmitt & Clapham, 2001), vocabulary knowledge grows progressively from the simple, most frequent words towards the least frequent, more specialized words.

Different researchers have provided figures regarding the number of words we need to know to perform certain tasks. For speaking, Adolphs and Schmitt (2004) claim that a vocabulary of around 2,000 words is required for a basic conversation, but according to other vocabulary specialists, if students want to obtain an appropriate comprehension of a text, they need a passive vocabulary of at least 3,000 word families to understand 95% of a text (Laufer, 1989, 1992); or from 8,000 to 9,000 word families for a 98% text coverage (Hu & Nation, 2000; Nation, 2006). As for listening, Nation (2006) claims that lexis of 6,000 to 7,000 word families is required to comprehend oral texts successfully. What this indicates is that, as Schmitt (2010: 4) contends, “vocabulary has strong relationships with the language skills”. The sample chosen here represents a social group at a very significant stage in the educational process, the result of a learning process encompassing their primary and secondary education.

One of the purposes behind determining how much vocabulary students know is to use this information to identify the relation between vocabulary size and the effect that this knowledge of vocabulary exerts on language use. In the Costa Rican educational system, students are required to take a series of tests to determine whether they can continue their education at the university level. One test is that of English reading comprehension, reading being the most important element in the country’s curriculum. Hirsh and Nation (1992: 689) claim: “[k]nowledge of the vocabulary in a text is one of the main factors that affect reading”. This use of *passive* vocabulary that students implement for reading is only one of the dimensions in which vocabulary is used daily; however, as discussed above, reading marks the way in the Costa Rican curriculum. Along these same lines, Nation (1993: 132) insists, “[...] if learners do not develop a sufficiently large vocabulary then skill in language use and the knowledge that comes from that skill will not develop”.

With that in mind, the underlying purpose of researchers and educators in examining vocabulary progress should be directed towards finding ways of accompanying students in the process of developing the skills required for using the language effectively. Vocabulary represents much more than just a number of concepts that students know. For Read (2000: 14), “vocabulary is not just a set of linguistic units but also an attribute of individual language learners, in the form of vocabulary knowledge and the ability to access that knowledge for communicative purposes”. Researchers should find ways that would eventually facilitate and

ensure real communication for students. To achieve that, we must find out what students know in terms of vocabulary, so that the necessary adjustments can be made in both curricula and classrooms. Read (2007: 121) states, “[w]e need to complement discrete vocabulary tests with embedded measures of the learner’s ability to handle lexical items in context. Traditionally, context has been conceived in linguistic terms as a sentence or larger co-text in which a vocabulary item occurs”. This is exactly the task required of students in the test they take upon completion of high school. The above ideas acknowledge the intertwined connection between vocabulary and L2 acquisition, and in the particular case of the Costa Rican students, between vocabulary knowledge and reading. The role of vocabulary is key to benefit or impair students’ progress in the reading process; hence the importance of determining the impact that vocabulary knowledge may have on students’ performance on the pre-university test they are required to take.

As seen above, the area of vocabulary studies is only recently developing more systematically, and the breadth of studies is not as vast as it is for other language areas. Regarding vocabulary measures, Terrazas Gallego and Agustín Llach (2009) offer a review of over 30 studies dealing with English native and non-native speakers’ vocabulary size. In this review, the participants range from university to primary school age students. Certain studies may serve as referents in the case that concerns us. Table 1 summarizes the studies in Terrazas Gallego and Agustín Llach (2009), as well as Canga Alonso (2013), in which participants are reported to be high-school students whose vocabulary is measured through the Vocabulary Levels Test (VLT):

| STUDY | REPORTED VOCABULARY SIZE | HOURS OF INSTRUCTION | STUDENTS’ L1 | LEARNING CONTEXT: ESL/EFL EDUCATIONAL LEVEL |
|-------------------------|--------------------------|----------------------|--------------|--|
| Qian (2002)* | 7,224 | --- | Korean | EFL/Secondary School (intermediate level and beyond) |
| Qian (2002)* | 6,663 | --- | Chinese | EFL/Secondary School (intermediate level and beyond) |
| Laufer (1998)* | 3,500 | 1,500 | Hebrew | EFL/Secondary School |
| Canga Alonso (2013) | 935 | 1,049 | Spanish | EFL/Secondary School |
| López-Mezquita (2005)** | 941 | 1,049 | Spanish | EFL/Secondary School |

Table 1. Previous studies as reported in *Terrazas Gallego and Agustín Llach (2009) and **Canga Alonso (2013).

As observed in Table 1, Terrazas Gallego and Agustín Llach (2009) refer to three studies conducted with secondary school students. They cite Qian (2002), who, according to

Terrazas Gallego and Agustín Llach (2009), determines a vocabulary size of 7,224 for Korean secondary students and 6,663 for Chinese secondary students. They also discuss a study by Laufer (1998) in which the passive vocabulary size of a group of Hebraic secondary students is 3,500 words. For the Asian students no mention is made of the number of hours of instruction, while for the Hebraic students they cite 1,500 hours. The last two studies will be addressed in the discussion section. Canga Alonso (2013: 69), on the other hand, establishes a considerably lower mean of 935 words for a group of secondary, Spanish native-speaker students who have received 1,049 hours of instruction. Canga Alonso (2013: 67) also refers to a study by López Mezquita (2005), who determines a vocabulary size of 941 for a group of Spanish secondary students who have received 1,049 hours of instruction. No studies of this nature have ever been implemented in the Costa Rican context. The present study will provide us with essential information for this undeveloped area of study in Costa Rica.

As concerns gender studies in connection with second language acquisition, literature is also limited. Some studies deal with gender and strategy use (Jiménez Catalán, 2003; Lee, 2007), attitude and interests for learning a second language (Baker & MacIntyre, 2000), or gender and self-efficacy in learning (Hampton & Mason, 2003). As for gender and vocabulary acquisition, in particular, studies are also very scarce and thus inconclusive. Grace (2000) analyzed retention of lexicon and the use given to translation practices using an educational computer program. Subjects' receptive vocabulary was evaluated after students had received instruction through a computer program that offered either translation of vocabulary to their L1 or an explanation of vocabulary through the L2 language. Grace found no significant differences between males and females in the translation or non-translation groups for the type of sentence translation elicited by students, nor for the total amount of time that each gender group invested in the given translation option. Grace (2000) claimed that both gender groups benefited equally from a lesson taught completely through the L2 language. Interestingly, she also noted that, regardless of gender, vocabulary test scores were even better for the groups with access to the L1.

In the following three studies, vocabulary was also measured through means different from the VLT. In that of Lynn, Fergusson and Horwood (2005), interesting gender differences were found in a series of tests evaluating various areas of language and other skills. When the task involved coding ability, for example, girls had better results than boys did, whereas boys outscored girls on tests evaluating vocabulary and information. In another vocabulary investigation, Lin and Wu (2003: 9–10) compared differences in gender results in high stake examinations of a group of over 4,400 adults. Their study reports that when items are bound according to their differential functions, statistically significant results could be identified: girls performed better in listening comprehension items, and boys outperformed girls in grammar, vocabulary and cloze items, albeit slightly. Heras and Lasagabaster (2015), on the other hand, found no significant gender differences in terms of vocabulary measures despite the initially seeming advantage of female students.

For vocabulary studies that used the VLT, a couple of cases can be mentioned. Canga Alonso (2013) reported statistically significant, higher receptive vocabulary measures for boys in total word counts of 1,028 as opposed to 854 words for girls. Contrariwise, also based on VLT results, different outcomes (in terms of male vocabulary advantage) were found by Jiménez Catalán and Terrazas Gallego (2005) in younger learners, whereby the results showed that girls exhibit a certain advantage in vocabulary counts. The authors insist, however, that this difference does not appear to be significant. The current study thus aims to reduce the scarcity of information on this area, determining whether differences exist and whether these could result in direct modifications of classroom practices.

Turning to the scope of the present study, additional considerations are necessary. For the sake of clarity and considering the breadth of definitions of terms such as *Content based teaching* and *Foreign Language teaching*, these terms will be used here along the following lines: Baker (2011: 217–218) defines *Foreign Language teaching* as an educational system in which “most language majority school children take their education through their home language [...] Second (foreign) language lessons of around half an hour per day may constitute the sole ‘other’ language diet [...] This is distinct from teaching through the medium of a second language where curriculum content is the main focus rather than language learning [...]”. In this paper, whenever *Foreign Language (FL) school* is used, it refers to an institution characterized by the above definition. In Costa Rican mainstream education the foreign language is taught as one of the subjects in the curriculum. In the case of *Content based teaching*, Dupuy (2000: 212) defines it as “teaching a content area in the target language wherein students acquire both language and subject matter knowledge”. The author emphasizes features such as (1) form and not language being the focus of curriculum; and (2) instructors providing students with comprehensible input. According to Dupuy (2000), in content-based instruction an attempt is made to resemble first language acquisition. Through that focus on meaning rather than form, the language input is slightly above that of the students and aims to provide an opportunity for students to use the language meaningfully. This is the case of what is referred to below as a *Content based (CB) school*.

Practices such as content-based teaching invite us to consider more updated, advantageous methods such as CLIL. Mehisto, Marsh and Frigols (2008) describe CLIL as a teaching approach using an additional language to learn not only content but also language while also providing students with independence and a cooperative sense towards others. For authors such as Juan-Garau and Salazar-Noguera (2015: 3), CLIL describes teaching practices inside “an umbrella term broadly covering the central part of [the] continuum between content-driven and language-driven teaching approaches”. According to this description, we could include the content-based instruction practiced in Costa Rica as part of the CLIL methodology. However, Coyle, Hood and Marsh (2010) expect more from CLIL practices than what is usually found in traditional content-based instruction. For them, the four Cs—content, communication, cognition, and culture, which are CLIL’s cornerstones—

bring more to the knowledge students can finally gain. Coyle, Hood and Marsh (2010: 6) affirm that “[w]hat separated CLIL from some established approaches such as content-based language learning or forms of bilingual education is the planned pedagogic integration of contextualized content, cognition, communication and culture into the teaching and learning practice”. In the latter perspective, CLIL goes one step beyond the former idea conceived in content-based instruction. Different studies already point us toward CLIL methodology and the benefits it could have for students in the Costa Rican context. Of particular interest to the present study is Dalton-Puffer’s study (2008), which explains how receptive skills and vocabulary both benefit from CLIL practices. There is some controversy on the subject, though, which can be taken as a call for more research on the subject. Agustín Llach (2015) claims no vocabulary advantage for primary CLIL students over EFL learners, allegedly due to the young age of the participants. Agustín Llach and Canga Alonso (2016) contend significantly higher vocabulary scores over time and thus larger vocabulary sizes for CLIL over EFL primary students. Gierlinger and Wagner (2016) associate CLIL success in vocabulary learning to the frequency of input in this type of practice. The present study contributes some information to this ongoing discussion.

2. THE PRESENT STUDY

2.1. Premises and research questions

On the premise of the vocabulary theory addressed above, knowledge of VLT 2,000 is essential for different reasons. First, Nation (2001) describes high-frequency words contained in the 2,000 level as a determinant basis for language learning. Second, Nation (2006) insists that the 2,000 vocabulary word band, as opposed to upper word levels, requires classroom time and more directed attention; thus, knowing whether this is happening in the Costa Rican schools is deemed important. Third, Nation (1983, 1990) and Read (1988) contend that knowledge of the 2,000-word level band is essential to proceed with the acquisition of upper-level word bands. Fourth, Laufer and Nation (1999: 36), following Nation and Hwang (1995), describe the 2,000-word level as the “dividing line between high- and low-frequency words”; thus establishing the 2,000-level word band as the key for access to more specialized vocabulary. In that vein, this investigation aims to determine how many of those 2,000 key words are known by the students participating in this study, with the intention of defining vocabulary word counts and their possible impact on language use for the students in the sample. This leads then to the following research questions:

1. What is the vocabulary size of secondary students in two Costa Rican high schools, each of which follows different teaching practices?
2. Is there a difference in vocabulary size in different gender groups at these schools?

2.2. Participants

The subjects under study are 185 eleventh-year, secondary students who are learners of English as a foreign language in two middle-class high schools in Heredia, the third most populated province in Costa Rica. There are 101 girls and 85 boys; 55 of these students attend a school where content-based instruction is implemented, and 130 study at a school where they take English as a compulsory course, in a regular, mainstream Foreign Language setting. The sample is homogenous in terms of L1, age and social profile of the area where the schools are located. Spanish is the first language for all the students. None have English native-speaking parents, nor have any lived in an English-speaking country. In each of the schools, participants had been taught English by the same teachers throughout their high school studies. The average age of the students is 16.6 years. Both schools also follow the same national curriculum for the foreign language courses. Regarding the total number of hours of instruction at high school level, the CB school students have received approximately 1,368 hours, and the mainstream EFL school students have received approximately 1,140 hours of class. This translates into 304 hours of English instruction per year during 7th, 8th, and 9th grade, plus 228 hours of English instruction in 10th grade for the FL school; and 361 hours of instruction in 7th, 8th, and 9th grade, plus 285 hours of instruction in 10th grade for the CB school. The hours of instruction mentioned above for the CB school include 3 hours a week in a one-subject course that varies from one level to the next: Ecology, Social Studies, and Biology in 7th, 8th, and 9th school years, respectively.

2.3. Instruments

The VLT, 2,000 band Version 2 (Schmitt, Schmitt & Clapham, 2001), is used to determine students' vocabulary size. In this test the participants must complete a total of 30 items. For each one, they match a word to its corresponding definition (there are 30 extra words for which no definition is given). This test provides the information required to profile students in terms of their estimated vocabulary size. According to Schmitt, Schmitt and Clapham (2001), the 2,000 level contains words from the first 1,000 words and from the second 2,000 words in a ratio close to 1:1 (28 from the first 1,000 frequency level and 32 from the second 1,000 level). Research has shown the VLT to be valid (see Beglar & Hunt, 1999; Cameron, 2002; Laufer, 1998; Schmitt, Schmitt & Clapham, 2001), reliable (Read, 1988; Schmitt, Schmitt & Clapham, 2001), and practical (Schmitt, Schmitt & Clapham, 2001). Through the use of this instrument, we try to determine how many words students can understand without the help of contextual cues.

2.4. Procedure

Data were collected in one session per test in each school, during the English class in all six sub-groups; that is, two 11th year groups in the CB high school and four 11th year groups in the FL high school. Students were allowed 15 minutes to complete the VLT, although many of them finished earlier. At the beginning of the VLT, both oral and written instructions were given in Spanish. Descriptive statistics were calculated and differences between schools and gender groups were compared by using SPSS Version 20 (IBM Corp, 2011).

2.5. Results

A description of the results is presented below for each of the schools participating in this study. First, a description of the total number of words in each school is provided, and, then, the distribution of words by gender is given for each institution.

2.5.1. Results in relation to word counts

Table 2 presents the first results for mean and standard deviation. When we analyze the mean from the perspective of the 30 items in the test we see that it is higher for the CB school; it is placed at 26.91 against 22.15 for the mainstream FL school. The standard deviation numbers also show a major concentration of the data for the CB school whereas the range of data for the FL schools shows greater dispersion. These results may be analyzed in greater depth with relation to the 2,000 words that are measured in the VLT 2,000. Following Nation's (1990: 78) formula "Vocabulary size = N correct answers multiplied by total N words in dictionary (the relevant word list) divided by N items in test", we convert the total of correct items into total word counts for each student in each school. Table 2 also shows that the CB school students achieve a mean of 1,793.94 words out of the 2,000 possible words and that the FL school students reach a mean of 1,475.85 words. Interestingly, we can also observe that although some students in both schools reach the maximum 2,000 words, the minimum of words is much lower (600) for students in the FL school, hence the higher spread of data expressed by the standard deviation observed for this school (349.7) when compared to that of the CB school (152.9), where the minimum number of words was 1,400:

| VLT 2,000 (n=185) | | | | |
|------------------------|-----------------|-----------|-----------------------|-----------|
| | ITEM COUNT DATA | | TOTAL WORD COUNT DATA | |
| | CB School | FL School | CB School | FL School |
| Total numbers | 30 | 30 | 2,000 | 2,000 |
| Mean | 26.91 | 22.15 | 1,793.94 | 1,475.85 |
| Max items/words | 30 | 30 | 2,000 | 2,000 |
| Min items/words | 21 | 10 | 1,400 | 600 |
| SD | 2.29 | 5.24 | 152.9 | 349.7 |

Table 2. Mean, standard deviations and word estimates for schools.

Figure 1 illustrates the range of distribution for both schools, and facilitates the comparison between them. While we can see that for the CB school the lowest score falls around 1,400 words, we can also observe that most of the distribution is clustered above 1,600 words and even more tightly clustered between 1,800 and 1,950 words. On the other hand, Figure 1 also shows that the range distribution for the FL school is much more spread along the vocabulary band. As Figure 1 shows, the distribution starts around 600-word range and unevenly moves up to 2,000 words. Although the figure shows a greater clustering of data starting around 1,200 and up to 2,000 words, the highest point in this portion of the data appears to cluster around 1,700 words:

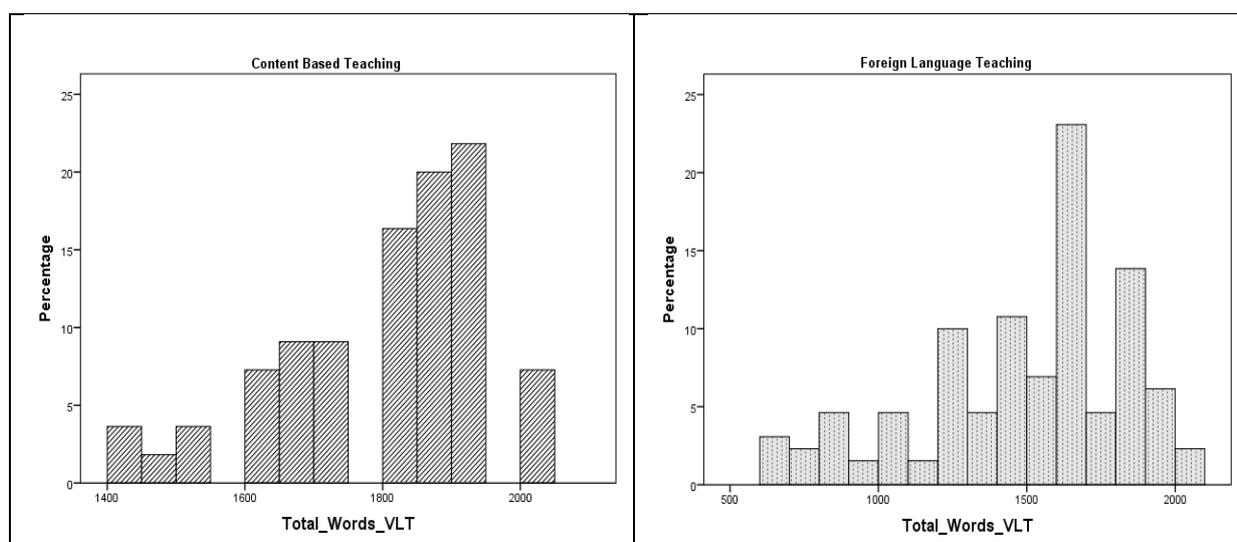


Figure 1. Percentages of distribution of words for CB school and FL school.

Cronbach's coefficient ordinal alpha was conducted to check for internal consistency reliability estimates and according to this test the CB school's alpha is 0.86 and the FL school's alpha is 0.94. In both cases the results can be interpreted as reliable. Although for the FL school the alpha seems high, according to Larson-Hall (2010), this result is normal given that the number of students in the FL school is larger than that of the CB school. Kolmogorov-Smirnov and Shapiro-Wilk were used to determine normality assumptions in our data, with the purpose of defining the nature of the differences between the results of both schools in the VLT (Table 3):

| | HIGH SCHOOL | KOLMOGOROV-SMIRNOV | | SHAPIRO-WILK | | | |
|--------------|-------------|--------------------|------|--------------|------|-----|------|
| | | gl | Sig. | gl | Sig. | | |
| Rec_Num_Wrds | CB school | .174 | 55 | .000 | .913 | 55 | .001 |
| | FL school | .139 | 130 | .000 | .932 | 130 | .000 |

Table 3. Parametric test for schools.

According to this test, the data follows an abnormal distribution which calls for non-parametric tests of mean comparison for two independent variables; using the median is more

appropriate in this case. The U Mann Whitney test was therefore conducted to measure inferential statistical differences between the groups. As seen in Table 4, this test reveals statistically significant differences in favor of the CB school, which shows a mean rank of 131.70 over a 76.63 for the FL school:

| | |
|----------------|------------------|
| n=185 | 2,000 VLT |
| Mann Whitney U | 1,446.500 |
| Wilcoxon W | 9,961.500 |
| Error | 331.762 |
| Z | -6.416 |
| p (two-tailed) | .000 |

Table 4. Results of inferential statistics for the 2K VLT.

2.5.2. Results in relation to gender

The results are also analyzed from a gender perspective. The data in this sample, presented in Table 5, provide evidence that the boys’ overall receptive vocabulary is slightly higher than that of girls’ in both schools. Although the average word-count difference is low (fewer than 100 words between boys and girls in each school), the data still evinces an advantage for boys in terms of total word counts. The standard deviation information is also interesting as it shows a slightly broader diffusion in the data of word counts in the girls’ group for the CB school and a larger dispersion in the results for boys rather than girls in the FL school:

| | CB SCHOOL | | FL SCHOOL | |
|---------------|-----------|----------|-----------|----------|
| | Boys | Girls | Boys | Girls |
| Mean* | 1,828.57 | 1,772.53 | 1,512.45 | 1,440.35 |
| Median | 1,867.00 | 1,800.00 | 1,600.00 | 1,533.20 |
| SD | 142.35 | 157.43 | 361.44 | 336.88 |

Table 5. Mean and standard deviation based on gender in both schools. [The maximum number of words in this test is 2,000].

Once again, normality assumptions are checked for each gender group at each school to establish the characteristics of the differences found. Kolmogorov-Smirnov and Shapiro-Wilk tests were used (see Table 6):

| SCHOOL | GENDER | KOLMOGOROV-SMIRNOV | | | SHAPIRO-WILK | | |
|--------|--------|--------------------|----|------|--------------|----|------|
| | | | gl | Sig. | | gl | Sig. |
| CB | Boys | .225 | 21 | .007 | .883 | 21 | .015 |
| | Girls | .1157 | 24 | .032 | .910 | 34 | .009 |
| FL | Boys | .158 | 64 | .000 | .920 | 64 | .000 |
| | Girls | .139 | 66 | .003 | .930 | 66 | .001 |

Table 6. Normality distribution test for gender word counts at each school.

Given the abnormal distribution of the data according to gender, inferential statistics are calculated with the U Mann Whitney test (see Table 7):

| | 2K VLT | | |
|-------------|--------|----|-------|
| HIGH SCHOOL | GENDER | N | RANK |
| CB | Girls | 34 | 25.76 |
| | Boys | 21 | 31.62 |
| FL | Girls | 66 | 60.51 |
| | Boys | 64 | 70.65 |

Table 7. Ranks for 2K (VLT) and Reading Test.

Based on the Mann Whitney U test, no statistically significant differences are found in either school or gender group. This means that the apparent difference in word counts that favored male students is not statistically significant in the end (see Table 8):

| | CB SCHOOL | FL SCHOOL |
|-----------------------|-----------|-----------|
| Mann Whitney U | 281.000 | 1,782.500 |
| Wilcoxon W | 876.000 | 3,993.500 |
| Error | 281.000 | 1,782.500 |
| Z | -1.333 | -1.538 |
| <i>p</i> (two-tailed) | .182 | .124 |

Table 8. Results of inferential statistics for the 2K VLT in gender in both schools.

3. DISCUSSION

3.1. Total word counts

After analyzing the results of the test and looking at the results of the statistical information above, we can now attempt to describe the implications of the vocabulary measures of the students in this study. First, in terms of the VLT results and considering that Schmitt, Schmitt and Clapham (2001: 67) speak of a criterion of mastery of 26 out of the 30 possible items per level, we can conclude that the students in CB school do master this level (mean = 26.91; 1,793.94 words). On the other hand, although the students from the FL school do not meet this criterion (mean = 22.15; 1,475.95 words), they show much higher word counts than students in other studies in which Spanish is also the native language. For example, Canga Alonso (2013: 69) estimated 935 words; Fernández Fontecha (2014: 27) 985 words; and López Mezquita (2005) 941 words for the secondary students in their studies. Although in these three studies the difference in number of hours is only slightly lower than for the FL Costa Rican school (Canga Alonso [2013], 1,049 hours; Fernández Fontecha [2014], 839 hours; and López Mezquita [2005], 1049 hours), the number of words is significantly lower if

we consider that the hour difference is not extreme. On the other hand, the disparity of the results between the present study and those conducted in Spain might be seen as an indication to support Terrazas Gallego and Agustín Llach's (2009) claim that there is a steady increment of passive vocabulary gain as students advance in school grades. As indicated above, participants in this study are 16.6 years old, while those in Canga Alonso (2013) and Fernández Fotencha (2014) are 15–16 and 13–14 years old, respectively. Different from the results of the Spanish studies, the numbers for the Costa Rican students turn out to be extremely low in comparison to those of section 1 above, where we cited 7,224 and 6,663 words in Qian (2002) and 3,500 words in Laufer (1998).

A closer look at what Qian (2002) has done provides interesting data. Although Terrazas Gallego and Agustín Llach (2009: 117) classified the participants in Qian's study as "secondary school (intermediate level and beyond)", the description of the participants in Qian (2002: 523) does not coincide: "[m]ost of the participants were either university graduates ([n=]104) or undergraduates ([n=]75). The rest were high school graduates ([n=]21) and master's degree holders ([n=]7)". This would appear to indicate that the participants are beyond secondary school level. Moreover, Qian (2002: 522) notes that his subjects are attending an "intensive ESL program at the University of Toronto", and that many of them entertain the possibility of pursuing further studies in a university in an English-speaking country. The facts that most of the population analyzed by Qian (2002) had a stronger educational background than the participants in my study, that they are immersed in the target language environment (albeit recently), and that most demonstrate a clear interest in learning the target language (to the point of traveling to Canada for an intensive ESL program) are all elements justifying the exceedingly high numbers of word counts that Qian's (2002) group of participants exhibit. As for Laufer (1998), the participants in her study share most of the same features as the participants in the present study. After analyzing her study, it is difficult to determine what causes such evident total word count differences. However, she did not use the 2,000 test band alone; it seems that students took several bands of the tests in her study. In addition, she calculated the results in a rather unconventional way. In her descriptions of the figures, she mentions a 4th thousand-level score that results from the average of the 3,000 and 5,000 levels. For the calculation she says that the sum of scores was multiplied by 5,000 and divided by 108, and she points out that her students were tested only on 72 items (Laufer, 1998: 269). Though it is difficult to affirm with certainty that this is the case, this different way of calculating the total scores may be responsible for the word count differences in that study. One may also argue that the fact that students took four different test-bands (2,000, 3,000, UWL and 5,000), and that the results for all of them are added up, would be reason enough to justify higher word counts in that case.

3.2. A small contribution to gender results

In terms of gender, the scarcity of studies dealing with vocabulary acquisition and gender makes it difficult to come to conclusive results. The data obtained here appear to go along the same lines as those presented by Canga Alonso (2013), where male students outperform their female counterparts. Nevertheless, as mentioned before, this difference is significant in his study, but not in the present one. Lynn, Fergusson and Horwood (2005) give boys significant advantage over girls in vocabulary measures, whose result is based on a *t* test for independent samples ($p < .05$, $d = .13$). Lin and Wu (2003) also found slightly superior statistically significant differences in favor of adult male participants on their grammar-vocabulary test. As can be abstracted from these studies, more research is required where vocabulary measures and gender connections are identified in order to come to more definitive results on the relation between the variables. The present study offers a small, new piece of evidence in relation to this less developed area of research. It is hoped that future studies will direct us to more conclusive findings.

3.3. Possible implications for the Costa Rican context

We now turn to the implications that these receptive vocabulary measures may have for students in the Costa Rican educational system. As seen above, we have to take into account that high school students must take an English-reading comprehension test as part of the series of tests at the end of high school. Experts may disagree on a definite number, but they have concluded that a vocabulary coverage between 95% (Laufer, 1989) and 98% (Nation, 2006) is required to read and comprehend texts, and that learners should know more than 2,000 words to read well. Considering the previous vocabulary estimates, we find that although not all the participants in this study exhibit that capacity just yet, the means are slightly more adequate for the CB school. Yet, a large number of students, especially from the FL school do not have this vocabulary size and thus lack the tools required to deal with authentic texts adequately. The advantage obtained by the CB school may be associated with the teaching practice that it embodies. In many respects, CB teaching maintains similarities with Content and Language Integrated Learning (CLIL) practices. Dalton-Puffer (2008: 5–6) describes a much stronger passive lexicon as one of the advantages of CLIL practices. She claims that receptive skills and vocabulary are two areas that are “favorably affected”. The teaching practice, thus, may have been beneficial for one of the groups in this study.

What is important, therefore, is that the corresponding authorities establish a course of action that would help make up for the difference in vocabulary measures in both types of schools. Nation (2014: 14) argues that “learning through input is feasible for learners of English as a foreign language if texts at the appropriate level are available”. Given that the basis of mainstream education in Costa Rica comes through reading, a very careful selection

of texts is called for. Schmitt (2007: 751) is very clear in stating: “[f]irst, a learner is unlikely to be able to acquire a wide vocabulary [...] through explicit learning alone [...] Second, learning a more achievable number of word families (2,000–5,000) can provide considerable rewards in the linguistic abilities they support. A significant percentage of this vocabulary can be realistically addressed explicitly over a period of time. Third, the most important words to target for explicit attention are the most frequent words in English”. Furthermore, Schmitt (2008: 354) affirms, “an effective vocabulary learning program needs to be principled, long-term, and one which recognizes the richness and scope of lexical knowledge”. Vocabulary must have a prominent role in our curricula; the research that has served as basis for this study suggests that this might not be the case in the Costa Rican context. This points to a combination of classroom practices that can open up the range of possibilities for learners to gain vocabulary. Keeping that in mind, teachers and planning authorities need to review what is happening in the foreign language classrooms today. This may be the first step that could be considered in mainstream schools to improve the results discussed above.

3.4. Limitations and future research

Future studies are required to overcome several limitations identified in this analysis. In the present study data was collected in one single instance. It would be ideal to measure the vocabulary level of participants as part of a longitudinal study to determine progress in vocabulary acquisition. Second, the number of instruction hours reported in this study is limited only to the secondary school history of the students; primary education is not considered. Future studies should try to account for primary English learning and its variation. Third, measures of VLT 3,000 would have been appropriate given the effect that that vocabulary size has on reading. Although theory predicts that to know the 3,000 word level students need to know the 2,000 level—and after studying the results of this study we know that the FL school students do not reach the 2,000 level and that the CB school students barely do so—, 3,000 word level counts could shed some light on a more general vocabulary count for the participants in this study. Upcoming research should also try to establish vocabulary word counts in connection to reading and other skills.

4. CONCLUSION

The results presented here embody a small part of a broader study dealing with vocabulary acquisition in two high schools in the province of Heredia, Costa Rica. The study responds to the need for research on vocabulary knowledge within the EFL Costa Rican context. Further

studies involving larger samples and considering students from the public sector, as well as the possibilities discussed above, are imperative. Several important initial findings have been presented here. First, the results show that students whose instruction follows CB practices exhibit a statistically significant advantage in relation to FL students. Second, the total word counts for the sample population in this study indicate a mean of 1,793.94 for CB school students and 1,475.85 for FL school students. Third, no statistically significant differences were found for gender in either one of the schools in the sample studied. The gender differences that favor male students persist only at the total word count level. Finally, considering that most 11th year students are part of mainstream (FL) education, educators must take action, given that this study places students from the FL school in a disadvantageous position. The results point to the need for the Costa Rican national board of education to look into CLIL practices and possibilities for their application in the Costa Rican context. Given that CLIL is associated with benefits in the European setting, consideration should be given to implementing CLIL in the national curriculum, rather than the methods used traditionally in these schools.

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. 32–52). Amsterdam: John Benjamins.
- Agustín Llach, M. P. (2015). The effects of the CLIL approach in young foreign language learners' lexical profiles. *International Journal of Bilingual Education and Bilingualism*, 20(5), 557–573.
- Agustín Llach, M. P. & Canga Alonso, A. (2016). Vocabulary growth in young CLIL and traditional EFL learners: Evidence from research and implications for education. *International Journal of Applied Linguistics*, 26(2), 211–227.
- Aitchison, J. (2012). *Words in the Mind: An Introduction to the Mental Lexicon*. West Sussex: Wiley-Blackwell.
- Baker, C. (2011). *Foundations of Bilingual Education and Bilingualism*. Bristol: Multilingual Matters.
- Baker, S. C. & MacIntyre, P. D. (2000). The role of gender and immersion in communication and second language orientations. *Language Learning*, 50(2), 311–341.
- Beglar, D. & Hunt, A. (1999). Revising and validating the 2000 Word Level and University Word Level Vocabulary Tests. *Language Testing*, 16(2), 131–162.
- Cameron, L. (2002). Measuring vocabulary size in English as an additional language. *Language Teaching Research*, 6(2), 145–173.
- Canga Alonso, A. (2013). Receptive vocabulary size of secondary Spanish EFL learners. *Revista de Lingüística y Lenguas Aplicadas*, 8, 66–75.

- Coyle, D., Hood, P. & Marsh, D. (2010). *Content and Language Integrated Learning*. Cambridge: Cambridge University Press.
- Dalton-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. 1–19). Heidelberg: Carl Winter.
- Dupuy, B. (2000). Content Based Instruction: Can it help ease the transition from beginning to advanced foreign language classes? *Foreign Language Annals*, 30(2), 205–223.
- Fernández Fontecha, A. (2014). Receptive vocabulary knowledge and motivation in CLIL and EFL. *Revista de Lingüística y Lenguas Aplicadas*, 9, 23–32.
- Gierlinger, E. M. & Wagner, T. A. (2016). The more the merrier – Revisiting CLIL-based vocabulary growth in secondary education. *LACLIL*, 9(1), 37–63.
- Grace, C. (2000). Vocabulary retention and access to translations for beginning language learners in CALL. *The Modern Language Journal*, 84(2), 214–224.
- Hampton, N. Z. & Mason, E. (2003). Learning disabilities, gender, sources of efficacy, self-efficacy beliefs, and academic achievement in high school students. *Journal of School Psychology*, 41, 101–112.
- Heaton, J. B. (1990). *Classroom Testing*. Essex: Longman.
- Heras, A. & Lasagabaster, D. (2015). The impact of CLIL on affective factors and vocabulary learning. *Language Teaching Research*, 19(1), 70–88.
- Hirsh, D. & Nation, P. (1992). What vocabulary size is needed to read simplified texts for pleasure? *Reading in a Foreign Language*, 8, 689–696.
- Hu, M. & Nation, I. S. P. (2000). Vocabulary density and reading comprehension. *Reading in a Foreign Language*, 23, 403–430.
- IBM Corp. (2011). *IBM SPSS Statistics for Windows, Version 20.0*. Armonk, NY: IBM Corp.
- Jiménez Catalán, R. M. (2003). Sex differences in L2 vocabulary learning strategies. *International Journal of Applied Linguistics*, 13(1), 54–77.
- Jiménez Catalán, R. M. & Terrazas Gallego, M. (2005). The receptive vocabulary of English foreign language young learners. *Journal of English Studies*, 5, 173–191.
- Juan-Garau, M. & Salazar-Noguera, J. (2015). Introduction: The Relevance of CLIL Education in Achieving Multilingualism on the Global Stage. In M. Juan-Garau & J. Salazar-Noguera (Eds.), *Content Based Language Learning in Multilingual Educational Environments* (pp. 1–10). London: Springer.
- Larson-Hall, J. (2010). *A Guide to Doing Statistics in Second Language Research Using SPSS*. New York, NY: Routledge.
- Laufer, B. (1989). What percentage of text-lexis is essential for comprehension? In C. Lauren & M. Nordman (Eds.), *Special Language: From Humans to Thinking Machines* (pp. 316–323). Clevedon: Multilingual Matters.
- Laufer, B. (1992). How much lexis is necessary for reading comprehension? In J. Pierre, L. Anaud & H. Béjoint (Eds.), *Vocabulary and Applied Linguistics* (pp. 126–132). London: Macmillan.
- Laufer, B. (1998). The development of passive and active vocabulary in a second language: Same or different? *Applied Linguistics*, 19(2), 255–271.

- Laufer, B. & Nation, P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing*, 16(1), 33–51.
- Lee, S. (2007). Vocabulary Learning Strategies of Korean University Students: Strategy Use, Vocabulary Size, and Gender. *English Teaching*, 62(1), 149–168.
- Lin, J., & Wu, F. (2003). Differential Performance by Gender in Foreign Language Testing. Paper presented at *The Annual Meeting of the National Council on Measurement in Education* (NCME). Chicago, Illinois, April 2–24. Retrieved January 13, 2017 from <http://files.eric.ed.gov/fulltext/ED478206.pdf>.
- López Mezquita, M. (2005). *La evaluación de la competencia léxica: tests de vocabulario. Su fiabilidad y validez*. Unpublished Doctoral Dissertation, University of Granada, Spain.
- Lynn, R., Fergusson, D. M. & Horwood, L. J. (2005). Sex differences on the WISC-R in New Zealand. *Personality and Individual Differences*, 39(1), 103–114.
- Meara, P. M. (2009). *Connected Words: Word Associations and Second Language Vocabulary Acquisition*. Amsterdam: John Benjamins.
- Mehisto, P., Marsh, D. & Frigols, M. J. (2008). *Uncovering CLIL: Content and Language Integrated Learning in Bilingual and Multilingual Education*. London: McMillan.
- Milton, J. (2010). The development of vocabulary breadth across the CEFR: A common basis for the elaboration of language syllabuses, curriculum guidelines, examinations, and textbooks across Europe. In I. Bartning, M. Martin & I. Vedder (Eds.), *Communicative Proficiency and Linguistic Development: Intersections Between SLA and Language Testing Research* (pp. 211–231). Rome: EUROSLA.
- Nation, I. S. P. (1983). Teaching and testing vocabulary. *Guidelines*, 5(1), 12–25.
- Nation, I. S. P. (1990). *Teaching and Learning Vocabulary*. New York, NY: Newbury House Publishers.
- Nation, I. S. P. (1993). Vocabulary size, growth and use. In R. Schreuder & B. Weltens (Eds.), *The Bilingual Lexicon* (pp. 115–134). Amsterdam: John Benjamins.
- Nation, I. S. P. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, 63(1), 59–82.
- Nation, P. (2014). How much input do you need to learn the most frequent 9,000 words? *Reading in a Foreign Language*, 26(2), 1–16.
- Nation, P. & Hwang, K. (1995). Where would general service vocabulary stop and special purpose vocabulary begin? *System*, 23, 35–41.
- Nation, I. S. P. & Webb, S. (2011). *Researching and Analyzing Vocabulary*. Boston, MA: Heinle Cengage Learning.
- Pigada, M. & Schmitt, N. (2006). Vocabulary acquisition from extensive reading: A case study. *Reading in a Foreign Language*, 18, 1–28.
- Qian, D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning*, 52(3), 513–526.
- Read, J. (1988). Measuring the vocabulary knowledge of second language learners. *RELC Journal*, 19, 12–25.

- Read, J. A. S. (2000). *Assessing Vocabulary*. Cambridge: Cambridge University Press.
- Read, J. (2007). Second language vocabulary assessment: Current practices and new directions. *International Journal of English Studies*, 7(2), 105–125.
- Schmitt, N. (2000). *Vocabulary in Language Teaching*. Cambridge: Cambridge University Press.
- Schmitt, N. (2007). Current perspectives on vocabulary teaching and learning. In J. Cummins & C. Davison (Eds.), *The International Handbook of English Language Teaching. Volume 2*. (pp. 745–759). Berlin: Springer.
- Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research*, 12(3), 329–363.
- Schmitt, N. (2010). *Researching Vocabulary: A Vocabulary Research Manual*. Hampshire: Palgrave Macmillan.
- Schmitt, N., Schmitt, D. & Clapham, C. (2001). Developing and exploring the behavior of two new versions of the Vocabulary Level Test. *Language Testing*, 18(1), 55–88.
- 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.