

Are We Preparing Secondary Students for a Productive Use of Vocabulary in English as Their Second Language?

DAMARIS CASTRO GARCÍA
Universidad Nacional, Costa Rica

Received: 8 April 2017/ Accepted: 17 May 2017
ISSN: 1698-7467

ABSTRACT: This study examines the productive vocabulary profiles of Costa Rican high school students studying EFL under two different methodologies, Content Based Teaching (CBT) and Foreign Language Teaching (FLT). Following Laufer and Nation's (1999) idea of "controlled productive ability," the participants were tested using the Productive Vocabulary Levels Test (PVLTL). Statistically significant differences favor CBT students. After examining gender differences, the evidence also suggests that CBT could have a beneficial effect on students learning of the first 2,000 words of English, particularly for male students. Neither group, however, has fully mastered this word band yet.

Keywords: PVLTL, secondary, Costa Rica, productive vocabulary.

¿Estamos preparando a los estudiantes de secundaria para un uso productivo de vocabulario en inglés como su segunda lengua?

RESUMEN: Este estudio describe características del vocabulario productivo de estudiantes de secundaria costarricenses quienes estudian según las modalidades de Enseñanza por Contenidos o Enseñanza de Inglés como Lengua Extranjera. Siguiendo la idea de "habilidad productiva controlada" (Laufer y Nation (1999)), los estudiantes realizaron la Prueba de Nivel de Vocabulario Productivo. Los resultados muestran diferencias estadísticas significativas a favor de los estudiantes según el método de Enseñanza por Contenidos tanto en relación a niveles totales de palabras como en cuanto a género (particularmente para los hombres). Sin embargo, ninguno de los dos grupos domina completamente este nivel de vocabulario.

Palabras clave: PVLTL, secundaria, Costa Rica, vocabulario productivo.

1. INTRODUCTION

The study of vocabulary in second language learning has recently received an upswing in attention from researchers and other specialists. In the last twenty years, an increasing number of studies on the field have surfaced (i.e., Nation (1983, 1990, 2006 a, b), Laufer (1998), Read (2000), Schmitt (2010)). Laufer and Nation (1999:33-34) acknowledged this upsurge of vocabulary studies: "This increased interest can be interpreted to mean that there is considerable value in gaining knowledge about specific parts of language learners' proficiency because it can be used effectively for diagnostic, placement and curriculum-design purposes." Several other reasons have been mentioned about the importance of vocabulary studies. Nation (1993) insists that an extensive vocabulary is critical for students to use

language skillfully and for them to obtain the knowledge that derives from this use. This shows the importance of vocabulary development as a basis for the proper progression of language skills, in particular, and cognitive development in general. 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.” From this, we can gather that vocabulary no longer stands as a separate component of language attributes, but rather it is acknowledged as a cornerstone for building and developing language capacities instead.

Time has progressively lent a more prominent role to vocabulary than it enjoyed previously. Schmitt (2010), for example, also points to the strong link existing between vocabulary and language skills. Along these lines, Coombe (2011) states that vocabulary has gained importance as a skill on its own rather than functioning as an aid to the four main language skills. This new perspective on vocabulary has granted it the deserved attention that has resulted in more in-depth studies on the field.

Seminal work by Nation (1983, 1990) has served as a guide for vocabulary studies. Of particular significance for this paper is his determination in establishing vocabulary measures for teaching practices and program designs and his contribution to the development of the Vocabulary Levels Test (VLT), as described in Nation (1983, 1990) and Schmitt et al. (2001), and the subsequent Productive Vocabulary Levels Test (PVL), as discussed in Laufer & Nation (1999). Both versions include a set of tests divided into word bands according to the frequency of occurrence of words in the language: 2,000, 3,000, 5,000 Academic Word List (AWL), and 10,000 word bands.

In Nation’s word bands, the first 2,000 word list is made up of ‘high frequency vocabulary’ (Nation & Hwang, 1995), and these are the words that students should aim to learn first. Nation and Chung (2009), as well as Laufer & Nation (1999), discuss the importance of learning vocabulary according to its frequency and they insist on giving special attention to the first 2,000 words, not only because these represent 80 to 90% of running words in texts but also because knowing them allows learners to reap greater benefits. The 2,000 word band has also been identified as the dividing line between high and low frequency vocabulary (Laufer & Nation, 1999).

Nation (2006a) argues that the first 2,000 words of English represent the basic words of the language, and many are function words. Read (1988) also notes 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. For Read, and for Schmitt et al. (2001), vocabulary knowledge grows progressively from the simple, most frequent words towards the least frequent, more specialized words. Moreover, Webb and Nation (2011) point to how a distinction between different levels of proficiency can be attained through vocabulary measures. In that vein, this investigation aims to determine how many of those 2,000 key words are known by the participating students.

In sum, the above information suggests—following Read (2000:1)—that words can be considered “the basic building blocks of language, the units of meaning from which larger structures such as sentences, paragraphs, and whole texts are formed.” What interests us here is to determine how solid these building blocks are for the participants in the present study, specifically concerning their productive vocabulary. This, in turn, can provide us with information about what is currently happening at secondary schools like those in the

sample. According to Nation (2006b), knowledge of a word entails knowing aspects of its 'form, meaning, and use'; through the PVLТ, a controlled-productive activity, we attempt to determine the learners' potential use of these words in communicative contexts. Finally, we assume here the notion presented by Heaton (1990) regarding active or productive vocabulary as that which students ought to use. The PVLТ will require students to use the words they know to complete meaningful ideas in context.

2. PRODUCTIVE VOCABULARY

The ultimate intention of learning a language is to use that language for communication. To achieve that in an L2, learners require enough vocabulary to convey the meaning they wish to express. The number of words in our vocabulary, learned during the SLA process, would (ideally) increase gradually. Laufer (1998) comments that despite evidence of vocabulary growth in L2 learners, the numbers are radically lower when compared to native speakers' word counts. Moreover, Laufer (1998) insists that word knowledge advances from shallow to deep as students' learning progresses. This process takes time and effort and develops differently for different individuals under different conditions. Productive vocabulary has proven to be fundamental since early ages.

Vagh et al. (2009), for example, provide evidence for faster growing, higher-quantity, productive vocabulary in monolingual rather than bilingual children up to 36 months old. This shows how specific conditions may affect productive vocabulary development and also calls for consideration of the consequences of these differences. Likewise, Bleses et al. (2016) provide further rationale on how children's early productive vocabulary measures can predict future educational outcomes in decoding and reading comprehension abilities. They establish a link between early vocabulary measures and language development, and make a meaningful association between language and literacy (also present, to a lesser extent, for mathematical skills). Bleses et al. (2016) propose using oral skills, for which vocabulary plays a key role, as a strong basis for reading. These studies signal the long-term significance and impact of vocabulary on the educational future of monolingual and bilingual students. They are additional reasons to justify the need for additional productive vocabulary research. These findings from younger learners also extend to learners of other age groups.

In SLA studies, Laufer (1998:256) acknowledges that the process of vocabulary learning moves from receptive to productive expertise and that receptive vocabulary becomes larger. Any language learner would agree that both types are, indeed, essential during the learning process. By establishing the size of vocabulary (receptive and productive), teachers and program designers can trace the progress of students in any given language program. Laufer and Nation (1999:33) insist on the importance of determining learners' vocabulary knowledge to design the vocabulary component of teaching programs. As seen above, most studies on vocabulary measures have been guided and developed in connection to Paul Nation's (1983, 1990) and Schmitt et al.'s (2001) VLT, as well as Laufer and Nation's (1999) subsequent PVLТ. The latter is not found in academic research studies as frequently as the former. The same occurs with productive vocabulary, which has been researched less than receptive vocabulary. Thus, the importance of studies such as this one.

Laufer and Nation (1999:36) insist that the ability to produce vocabulary “implies degrees of knowledge” and claim that students can use vocabulary forms in some educational situations (i.e., when required by the teacher) and not in others (i.e., free writing). Laufer and Nation (1999:37) establish a distinction between “the ability to use a word at one’s free will as ‘free productive ability’ [and] ‘controlled productive ability’ for the ability to use a word when compelled to do so by a teacher or researcher.” The PVLTL measures the latter. Controlled productive ability is of particular importance because we can predict (given the idea of progressive degrees of knowledge) that students should be able to use vocabulary in these controlled settings before they can actually use it in activities with a freer context requiring deeper vocabulary knowledge such as writing.

Table 1 summarizes the only studies the author of the present paper has found on productive vocabulary measures where the PVLTL has been used. Four different reference groups are shown; the first two come from Laufer (1998). They are presented here separately with the test performance, hours of instruction and age. For the first two groups, Hebrew is the L1 and for the second two groups Spanish is their L1. All groups are studying English as a Second Language. Two important differences are observed: the number of words for the Israeli students is over double that of the Spanish students and some difference is found in hours of instruction.

Table 1. Previous Studies

Study	Participants	Hours of Instruction	Vocabulary Size
Laufer (1998)	Israeli, high school, 16 years old	1,080	1,700 words
Laufer (1998)	Israeli, high school, 17 years old	1,260	2,550 words
Moreno Espinoza (2010a) in Canga Alonso & Arribas García (2014)	Spanish, high school (age not reported)	Not reported	645 words
Canga Alonso & Arribas García (2014)	Spanish, high school, 15-16 years old	1,049	644 words

2.1. SLA, Productive Vocabulary and Gender

SLA studies dealing with gender and vocabulary are scarce. They could be classified into three groups. The first includes those that look at gender with relation to different aspects of SLA. Sunderland (2010) has acknowledged, “In SLA [...] sex (usually rendered ‘politely’ as gender) as an independent variable has been investigated less than other potentially relevant variables.” Studies such as those of Agustín Llach (2010), whose instrument was a composition; Fernández Fontecha (2010), who used a lexical availability test; Jiménez Catalán (2010), who used two word tests, a composition and a cue word test; and Moreno Espinoza (2010b), who used a word association test, have explored various features of productive vocabulary in connection to gender. They concentrate on several aspects of vocabulary and

are mentioned here to provide access to gender-related information and productive vocabulary in a broader sense, not in reference to PVL. These studies present mixed results as no clear tendencies are identified. Agustín Llach (2010), exploring lexical creations in early secondary education, found no significant gender differences in students' lexical creations (despite finding females superior in words-per-composition production). Fernández Fontecha (2010) studied 'gender and motivation' for productive vocabulary and found that female students statistically outperformed their male counterparts in both variables. Jiménez Catalán (2010) determined significant differences favoring female students in a composition and a cue-word productive test. In addition, Moreno Espinoza (2010b) reports no statistically significant gender difference in tests of word association.

A second group of studies aims to establish gender differences that can be explained in connection with teaching contexts, associated with aspects such as motivation. Fernández Fontecha and Canga Alonso (2014) in a study on gender and motivation between CLIL and non-CLIL students found non-CLIL settings to be significantly more motivating for students of both genders, while male students in CLIL settings are more motivated than their female counterparts. Heras and Lasagabaster (2015) report a difference in motivation between CLIL and non-CLIL settings. These authors in their study on gender, vocabulary and motivation found that female students exhibit greater motivation than male students in certain aspects of motivation, although not to a significant degree. The same happens in terms of vocabulary, where female students also show numerical advantages in terms of vocabulary results but these are not statistically significant at the end. These two references help illustrate the current contrasting situation between CLIL and non-CLIL settings. Although Content Based Teaching is not seen as a CLIL practice for some authors (Lasagabaster and Sierra (2010), Dalton-Puffer and Smit (2013)), others consider CBT a form of CLIL (Ruiz de Zarobe (2010)). The results of this study will also contribute to offer more information regarding the different outcomes of these settings.

In a third group of studies, we may include those covering gender differences in connection to the VLT and the PVL. Regarding gender differences in productive vocabulary tests (and in PVL in particular), the scarcity of studies is evident. From the research mentioned above, Canga Alonso and Arribas García (2014:47) report a mean of 636 words for male students and 661 for females. Despite initial differences, Canga Alonso and Arribas García (2014) conclude that no statistically significant differences prevail. No gender-related information was obtained from the other studies mentioned in Table 1.

As Fernández Fontecha (2010:105) puts it: "[w]hether supported by evidence or merely driven by uninformed stereotypical opinion, languages have been traditionally labeled as a female subject." While some studies presented above support this idea, others do not. Sunderland (2010) credits the frequently mentioned mixed results to aspects such as age, language level, classroom-setting activities and individual contexts. She finds it 'unsurprising' that all those features render these mixed results. All in all, the scarcity of studies describing productive vocabulary size and gender is evident. The present paper aims to contribute information to academic research on both of these fields.

3. THE PRESENT STUDY

3.1. Research Questions

This study attempts to provide answers for the following questions:

1. What is the productive vocabulary profile of students in two Costa Rican high schools?
2. Is there a significant difference between the test scores from the school that follows Content Based Instruction and the one that follows EFL methodology?
3. Is there a significant difference between the test scores for the gender variable at these schools?

3.2. Participants

The subjects are a total of 180 fifth-year, secondary students, learners of English as a foreign language in two high schools in Heredia, Costa Rica (in Central America). There are 96 girls and 84 boys; 53 of these students attend a school where Content Based (CB) teaching is implemented (20 boys, 33 girls), and 127 study at a school where they take English as a compulsory course, in a regular, mainstream Foreign Language (FL) setting (64 boys and 63 girls). The average age of the students is 16.3 years for the CB school and 16.5 years for the Foreign Language students. 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. Both groups also live in the same geographic area, (the central zone of the province of Heredia) and follow the same national language policy for language courses. The CB school adds 3 hours a week in one subject-course that varies from one level to the next: Ecology, Social Studies, Biology, in 7th, 8th, 9th year respectively. The CB school students have received approximately 1,368 hours, and the mainstream students approximately 1,140 hours. While the FL school is semiprivate and the CB school is private, both groups enjoy a school environment where English learning receives strong support, evident through presence of international students, international field trips, and the possibility of participating in clubs and projects requiring English. None of the members of the groups in either school had special education needs.

For the sake of clarity, the following definitions should be kept in mind in relation to what Content Based Teaching and Foreign Language Teaching entail. 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 explanation will describe, henceforth, ‘Foreign Language (FL) school’ in the context under study. As for 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.” This is the case of CB students in the present analysis.

3.3. Instrument

The Productive Vocabulary Levels Test (PVLТ) (Laufer & Nation 1999) is used to determine the ‘controlled productive vocabulary’ of the participants; particularly, the parallel version 1 (Version C) of the PVLТ-2,000 band. The test consists of 18 items where students have to complete a word that fits the context of a given sentence. A number of letters are provided in each case with the idea of triggering the target word. For Laufer and Nation (1999), this provision of letters serves a two-fold purpose. First, it provides as many letters as are required to disambiguate the cue, thus eliminating other possible alternatives; second, it has the intention of preventing learners from providing words from a different frequency band. This test has been found to be reliable, valid and very practical to implement (Laufer & Nation, 1999:44).

Sample: 1. I’m glad we had this opp_____ to talk.

3.4. Procedure

Data were collected in one session per group in each school, during the English class in all six sub-groups (two groups from the CB school and four groups in FL school). At the beginning of the PVLТ, both Spanish oral and written instructions were given to explain what the students were being asked to do. A sample sentence was also provided. Students were allowed 15 minutes to complete the PVLТ. Regardless of the institutional setting, some students completed the test in less than 15 minutes while others needed the full 15 min.

3.5. Data Analysis

Following Laufer and Nation (1999), the answers were checked based on semantic correctness. This means that they were graded as correct or incorrect, and that minor spelling mistakes (i.e., “lovley” instead of “lovely”) and grammatical mistakes (i.e., “introduce” instead of “introduced”) were ignored. Descriptive statistics were calculated and differences between schools and gender groups were compared by using SPSS Version 20 (IBM Corp, 2011). During the analysis of the data, a Cronbach’s alpha reliability test was run and it reported internal consistency of .75 for the CB school and .85 for the FL school.

3.6. Results

3.6.1. Results for word counts in the PVLТ

Table 2 presents the results for correct items and total word counts. Following Nation’s (1990:78) formula, “Vocabulary size = N correct answers multiplied by the N words in dictionary (the relevant word list) divided by N items in the test,” a relation was established between the number of items (maximum 18) and the total number of words that these represent in the 2,000 word band. For the total of 18 items, the CB school reaches a mean of 13.19 while the FL school mean is 9.30. Although in both schools students reach the maximum possible; the minimum for the CB school is 8 whereas for the FL school it is

zero. This difference is also reflected in a much wider spread expressed in the SD of 4.284 for the FL school as opposed to 2.836 of the CB school.

Table 2. Descriptive statistics for item and word estimates

PVLТ 2,000 (n = 182 (53 CBT, 127 FLT))				
	Item count information		Total word count information	
	CB School	FL School	CB school	FL school
Number of items	18	18	2,000	2,000
Mean	13.19	9.30	1,467.33	1,046.80
Max	18	18	2,000	2,000
Min	8	0	889	0
SD	2.836	4.284	312.075	477.644

Regarding total word counts, the mean is 1,467.33 words for the CB and 1,046.80 for the FL school. This shows that students in both schools reach a maximum of 2,000 words. The minimum, however, is clearly different: it stands at 889 words for the CB school and zero for the FL school. In an analysis of distribution, the frequency of distribution for the CB school presents two clearly separate word-level size groups. The first group of learners in the lower levels ranges from 800 to 1,400 words with a peak between the 1,200 and 1,400 words; then the data present a marked halt. The second cluster of learners' vocabulary size grows from the 1,600 word level with the highest peak at the 1,700 to 1,800 word range. It then slopes down to the 2,000 word range. The distribution for the FL school appears more evenly dispersed. The word sizes steadily grow from zero to the highest peak of 1,000 words; they then descend to 2,000 words. Additional possibilities of interpretation exist, such as eliminating the zero grades considering them as outliers (as suggested by an anonymous reviewer). At this point, all results are included to offer a more complete picture of the findings.

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 PVLТ, the Kolmogorov-Smirnov test was used. It confirms that the data follow a non-normal distribution for the CB school ($p = .001$), while the FL school data follow a normal distribution ($p = .200$). The U Mann Whitney test was then implemented to determine differences between the two school groups and the p value (.000) shows that there are significant differences favoring the CB school. The test shows a mean rank of 124.38 for the CB school and 77.33 for the FL school.

3.6.2. Results in relation to gender

According to Table 3, the data in this sample population show evidence that the overall productive vocabulary of male students is higher than that of girls in both schools; even if only slightly higher for the FL school. Both male and female students achieve the maximum word scores in the CB school, whereas only boys reach that number in the FL school. The spread in distribution shows a wider range of distribution for girls in the CB school, while this is much wider for boys in the FL school. In general, overall productive vocabulary measures are higher for both male and female students in the CB school.

Table 3. Descriptive statistics for gender and PVLТ

	CB school		FL school	
	Male (N=20)	Female (N=33)	Male (N=64)	Female (N=63)
Total words	2,000	2,000	2,000	2,000
Mean	1,592.52	1,387.67	1,091.97	1,001.63
SD	286.336	305.336	515.293	436.160
Max	2,000	2,000	2,000	1,888
Min	1,000	889	0	111

In general, it could be concluded that the behavior, in terms of vocabulary outcomes, is quite different for each educational institution. While vocabulary counts seem to exhibit very similar behavior for both genders in the FL school, the differences between gender groups in the CB school are evident. A Kolmogorov-Smirnov test was used on the data to establish normality assumptions for each gender group in each school. The test determines that the data for the CB school follows a non-normal distribution (masculine, $p = .000$; feminine, $p = .010$) while the data for the FL school follows a normal distribution (masculine, $p = .200$; feminine, $p = .200$). This calls for different statistical tests to explain the quality of the differences in each set of data. Given that the data for the CB school follow a non-normal distribution, the U Mann Whitney test was applied. The test shows a mean rank of 34.05 that favors male students and a mean rank of 23.33 for female students. It, thus, confirms that the distribution is not the same for the two genders and shows statistically significant results that favor male students in the PVLТ in CB school with a p value of .014.

Since the FL school data follow a normal distribution, the Independent sample t -test is used to determine the nature of those differences. The test shows that there are no significant differences between male and female students in this sample population of the FL school. The male student group ($N = 64$) obtained productive vocabulary level scores with a mean of 54.66 (SD 25.354), and female sub-sample, with a mean of 50.13 (SD 21.811). The results of the independent samples t -test are associated with a statistically non-significant effect: $t(126) = 1.084$, $p = .280$. This means that the numerical differences observed in the data do not translate into statistically significant differences for the gender variable at this school.

Additional tests were run, this time to compare gender results at inter-school level. As mentioned above, according to the Kolmogorov-Smirnov test, the distribution of the male CB school students follows a non-normal distribution ($p = .000$) while the FL male students' distribution follows a normal distribution ($p = .200$). Further examination using a U Mann-Whitney test shows that the numerically apparent differences between male students' scores are in fact statistically significant with a p value of .000 in favor of the CB male students. For the female students, the Kolmogorov-Smirnov test described the CB female distribution as non-normal ($p = .010$), whereas the FL female student group distribution was determined as normal ($p = .200$). The U Mann-Whitney test was implemented once more to determine the characteristics of the differences for total word counts in the PVLТ for the female gender across schools. According to this test, statistically significant differences ($p = .000$) favor CB female students, when we compare the two schools in the study sample.

4. DISCUSSION

The first information to be drawn from the data analysis above, considering the 15-16 correct-item requirement for satisfactory mastery established by Laufer and Nation (1999:41), is that neither of the groups in this study can be said to have a productive vocabulary size of 2,000 words. The CB school comes close at a mean of 13.19 but still does not make the cut. The FL school falls well below the requirement at a mean of only 9.30. While the differences between the two schools prove to be statistically significant, these differences do not guarantee full knowledge of the 2,000 word band for all students in the CB school group. These results, indeed, direct our attention to the FL school and its poor vocabulary attainment.

Comparisons of these results with those reported for previous studies in Table 1 reveal thought-provoking findings. If we compare our results with those in Laufer (1998), we discover that Laufer's (1998) groups reach much higher word counts despite having lower instruction hours. In her case, the 16-year-old group comes to 1,700 words (1,080 instruction hours), and the 17-year-old group reaches 2,550 (1,260 instruction hours). In the present study, the CB school reaches 1,467.33 words (1,368 instruction hours), and the FL school gets to 1,046.80 words (1,140 instruction hours). We could argue that this difference could be attributed to Laufer's (1998) participants taking all of the different word band tests whose results were added up to come to the grand total presented above. However, a closer inspection of Laufer's (1998:262) study reveals additional information. It shows that for the PVL 2,000 band alone, her 16-year-old group reaches an item mean of 11.83 and her 17-year-old group has a mean of 15 items. While only one of her groups reaches the suggested requirement, both of her groups are still well above the mean of the present study.

The other two possibilities for comparison presented in Table 1 tell quite a different story. When compared to the results found by Moreno Espinoza (2010a) and Canga Alonso and Arribas García (2014), the students in the present study appear to exhibit a clear advantage. Moreno Espinoza (2010) reports a word size of 645 while Canga Alonso and Arribas García (2014) speak of 644 words. For the latter study, both the age (15-16) and the total number of instruction hours (1,049) are a bit below those reported for the present study. The question remains whether this seemingly small difference in age and hours of instruction could account for the well-defined difference in total word counts between the Costa Rican and the Spanish students.

All other variables being equal, it would appear that the teaching method could have a different effect for male and female students in the CB school. The significant vocabulary advantage of the male students could be due to how students are dealing with the material covered through English instruction. A similar logic could apply for general vocabulary attainment between both schools. Inter-school comparisons reveal a statistically significant advantage for both male and female students in the CB school. This suggests that students do benefit more from CB teaching to learn vocabulary. As mentioned above, although the results show that students do not yet have full command of the 2,000 word band, CB students reflect a better mastery of this level. These results contrast with those of the literature

review section for Fernández Fontecha and Canga Alonso (2014) and Heras and Lasagabaster (2015). Although their studies initially focus on motivation, and that variable is not considered in the present study, we could compare the gender and language results of their studies. For Fernández Fontecha and Canga Alonso (2014), non-CLIL settings showed higher motivation, and we would expect this to translate into better vocabulary attainment for this group. However, the opposite is true in our study. It is the allegedly non-CLIL setting the one with lower vocabulary attainment. Heras and Lasagabaster (2015) found that CLIL contexts (allegedly the CBT in this study) help reduce gender differences in vocabulary learning. The present study, however, shows statistically significant differences favoring male students in the CBT setting. In an older study, Baker and MacIntyre (2000) found male students to exhibit limited positive attitudes toward learning French. In their study, involving immersion students, the non-immersion male students had worse results. We may speculate that the CBT setting may be exerting a positive effect on the male population in the present study.

As for the gender differences described in our study, they contribute slightly to the inconclusive results mentioned for gender studies and SLA. Given the scarcity of studies, this analysis can only be compared to that of Canga Alonso and Arribas García (2014). For those authors, the male and female participants exhibit non-statistical differences in their PVLТ results. In the present study, statistically significant differences were found between gender groups for the CB school, while the differences are non-significant in the FL school. It would appear that the treatment of vocabulary (or lack thereof) in the FL school produces the same effect for both genders. The gender similarity in productive ability in the test seems to suggest that it does. Further studies are required to determine the relevance of these conclusions.

These results call for reflection on the extent and quality of education that students are receiving in Costa Rica. It seems obvious that changes in the curriculum are necessary to provide students with better opportunities to produce language. Nation and Chung (2009), on the importance of providing students with the right opportunities for vocabulary learning, discuss the implementation of learning activities that rely on a set of equally balanced four strands of meaning (Nation 1996, 2007). These four strands are *meaning-focused input*, *meaning-focused output*, *language-focused learning*, and *fluency development* (Nation 2007). I would like to refer here to the *meaning-focused output* strand as one way to enable students to improve their productive vocabulary levels. Nation and Chung (2009) describe this strand as achievable through speaking and writing practices, and insist that if vocabulary were used to generate language creatively, this vocabulary would become strengthened in the students' memory. They suggest implementing activities such as retelling, rewriting or group negotiating tasks as prompts of this strand. While all four strands are of serious importance, this is one that could be said to be completely absent from the classroom contexts under study. Due to the constraints placed by the obligatory tests implemented by the Costa Rican Ministry of Education, the curriculum in Costa Rica relies heavily on reading. This is clearly to the detriment of students' productive abilities. Given the importance that knowing the first 2,000 words exerts on language development in general, a more productive orientation of the curriculum would be called for (such as that provided by CLIL, for example); that is, if we want our students to be able to use the language for real communication purposes.

5. CONCLUSION

The results of the present study indicate that the productive vocabulary profile of students in Costa Rica is limited and requires direct attention. Although neither school masters the 2,000 word level, statistically significant results in vocabulary favor the CBT practice over the FLT methodology. As for gender differences, male students show significantly richer vocabulary in the CB, while no gender differences are apparent in the FL school. We must analyze the incidence that these results have on much of the Costa Rican population, as CBT is limited to some private schools only. When learning a second language, students usually achieve a feeling of progress when they can find the words they need to express their ideas. Having access to the language that helps them convey their thoughts certainly gives learners a more concrete substantiation of their growth in the language acquisition process. While this idea of progression provides important motivation to actively continue in the process, the discouragement felt as a result of not knowing basic vocabulary can be detrimental. The study shows that, in their last year in high school, the participants in the study have not reached the minimum vocabulary levels required to know the high frequency, basic, productive, 2,000 band vocabulary. The numbers are even more alarming for the FL school. They reflect a sample of the mainstream education most of the Costa Rican teenage population receives. The educational system in Costa Rica requires attention and improvement. However, to provide a precise answer to the question presented in the title, the evidence appears to suggest that while some efforts are made to improve language teaching in Costa Rica, they are not enough to conclude that we are actually preparing our students to use English in real communication evident in language production. If “controlled productive ability” is the preceding stage to free productive ability as Laufer and Nation (1999) suggest, we still have much work ahead to pave the way to communication.

6. REFERENCES

- Agustín Llach, M.P. (2010). “Exploring the role of gender in lexical creations”, in R. M. Jiménez Catalán (ed.), *Gender Perspectives on Vocabulary in Foreign and Second Languages*. London: Palgrave, 74-92.
- Baker, C. (2011). *Foundations of Bilingual Education and Bilingualism*. Bristol: Multilingual Matters.
- Baker, S. C., and MacIntyre, P.D. (2000). “The role of gender and immersion and communication and second language orientations”, in *Language Learning*, 50: 311-341.
- Bleses, D., Makransky, G., Dale, P., Højen, A., and Aktürk Ari, B. (2016). “Early productive vocabulary predicts academic achievement 10 years later”, in *Applied Psycholinguistics*, 37: 1461-76.
- Canga Alonso, A. (2014). “Productive vocabulary knowledge of Spanish EFL learners”. *Revista Electrónica de Lingüística Aplicada 1*: 39-56, available from: <http://dialnet.unirioja.es/descarga/articulo/5031517.pdf>, accessed 19 January, 2017.
- Coombe, C. (2011). “Assessing vocabulary in the language classroom”, in *Malaysian English Language Teaching Association*, 111-124.
- Dupuy, B. (2000). “Content based instruction: Can it help ease the transition from beginning to advanced foreign language classes?”, in *Foreign Language Annals*, 30, 2: 205-23.

- Fernández Fontecha, A. (2010). "Gender and motivation in EFL vocabulary production", in R. M. Jiménez Catalán (ed.), *Gender Perspectives on Vocabulary in Foreign and Second Languages*. London: Palgrave, 93-116.
- Fernández Fontecha, A., and Canga, Alonso, A. (2014). "A preliminary study on motivation and gender in CLIL and non-CLIL types of instruction", in *International Journal of English Studies*, 14, 1: 21-36
- Heaton, J.B. (1990). *Classroom Testing*. Essex: Longman.
- Heras, A., and Lasagabaster, D. (2015). "The impact of CLIL on affective factors and vocabulary learning", in *Language Teaching Research*, 19, 1: 70-88.
- IBM Corp. (2011). *IBM SPSS Statistics for Windows, Version 20.0*. Armonk, NY: IBM Corp.
- Jiménez Catalán R.M. (2010). "Gender tendencies in EFL across vocabulary tests", in R. M. Jiménez Catalán (ed.), *Gender Perspectives on Vocabulary in Foreign and Second Languages*. London: Palgrave, 117-38.
- Laufer, B. (1998). "The development of passive and active vocabulary in a second language: Same or different?", in *Applied Linguistics*, 19, 2: 255-71.
- Laufer, B. and Nation, P. (1999). "A vocabulary-size test of controlled productive ability", in *Language Testing*, 16, 1: 33-51.
- Moreno Espinoza, S. (2010a). *Evaluación del vocabulario en redacciones escritas por aprendices del inglés como L2 a través de tecnología*. Unpublished PhD Dissertation. Universidad de la Rioja.
- Moreno Espinoza S. (2010b). "Boys' and girls' L2 word associations", in R. M. Jiménez Catalán (ed.), *Gender Perspectives on Vocabulary in Foreign and Second Languages*. London: Palgrave, 139-163.
- Nation, I.S.P. (1983). "Teaching and testing vocabulary", in *Guidelines*, 5, 1: 12-25.
- Nation, I.S.P. (1990). *Teaching and Learning Vocabulary*. New York: Newbury House.
- Nation, I.S.P. (1993). "Vocabulary size, growth, and use", in R. Schreuder & B. Weltens (eds.), *The Bilingual Lexicon*. Amsterdam: John Benjamins, 115-135.
- Nation, P. and Hwang K. (1995). "Where would general service vocabulary stop and special purpose vocabulary begin?", in *System*, 23: 35-41.
- Nation, P. (1996). "The four strands of a language course", in *TESOL in Context*, 6, 2: 7-12.
- Nation, I.S.P. (2006a). "Language Education-Vocabulary" in *Language Education: Teacher Preparation*. Oxford: Elsevier, 494-99.
- Nation, I.S.P. (2006b). "Vocabulary: Second language", in K. Brown (ed.), *Encyclopaedia of Language and Linguistics*. Oxford: Elsevier, 448-54.
- Nation, I.S.P. (2007). "The four strands", in *Innovation in Language Learning and Teaching*, 1, 1: 1-12.
- Nation, I.S.P. and Chung, T. (2009). "Teaching and testing vocabulary", in M. Long and C. Doughty (eds.), *The Handbook of Language Teaching*. Oxford: Blackwell, 543-59.
- Read, J. (1988). "Measuring the vocabulary knowledge of second language learners", in *RELC Journal*, 19: 12-25.
- Read, J.A.S. (2000). *Assessing Vocabulary*. Cambridge: Cambridge University Press.
- Schmitt, N., Schmitt, D. and Clapham, C. (2001). "Developing and exploring the behavior of two new versions of the Vocabulary Level Test", in *Language Testing*, 18, 1: 55-88.
- Schmitt, N. (2010). *Researching Vocabulary: A Vocabulary Research Manual*. Hampshire: Palgrave Macmillan.
- 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*. London: Palgrave, 1-19.

- Vagh, S.B., Pan, B.A., and Mancilla-Martinez J. (2009). "Measuring growth in bilingual and monolingual children's English productive vocabulary development: The Utility of Combining Parent and Teacher Report", in *Child Development*, 80, 5: 1545-63.
- Webb, S. and Nation, I.S.P. (2011). "Evaluating the vocabulary load of a written text", in *TESOLANZ Journal*, 16: 1-10.

APPENDIX. The Productive Vocabulary Levels Test (PVLТ) (Laufer & Nation 1999)**Estudiante:** _____

Colegio: _____

Grupo: _____

Tiempo: 15 minutos.

Prueba de Vocabulario Productivo (Laufer & Nation, 1999)

Complete las palabras que están subrayadas. Cuentan con un ejemplo completado para que vean lo que deben hacer.

Ejemplo:

Oración dada:

He was riding a bic_____.

Su respuesta:He was riding a bicycle.**Productive Vocabulary Test Level 2,000**

1. I'm glad we had this opp_____ to talk.
2. There are a doz_____ eggs in the basket.
3. Every working person must pay income t_____.
4. The pirate buried the trea_____ in a desert island.
5. Her beauty and cha_____ had a powerful effect on men.
6. La_____ of rain led to a shortage of water in the city.
7. He takes cr_____ and sugar in his coffee.
8. The rich man died and left all his we_____ to his son.
9. Pup_____ must hand in their papers by the end of the week.
10. This sweater is too tight. It needs to be stret_____.
11. Ann intro_____ her boyfriend to her mother.
12. Teenagers often adm_____ and worship pop singers.
13. If you blow up that balloon any more it will bur_____.
14. In order to be accepted into the university, he had to impr_____ his grades.
15. The telegram was deli_____ two hours after it had been sent.
16. The differences were so sl_____ that they went unnoticed.
17. The dress you're wearing is lov_____.
18. He wasn't very popu_____ when he was a teenager, but he has many friends now.