University Students Use of Academic Vocabulary in the BA in English and English Teaching

William Charpentier Jiménez
Escuela de Lenguas Modernas
Universidad de Costa Rica

Abstract

This article examines the amount of academic vocabulary used by fourth year students in a B.A. in English and B.A. English Teaching. The writer explains how academic vocabulary can be measured and the benefits it provides. This article contains a brief analysis of academic vocabulary used in 46 final research papers. Based on these results, the writer explains that, although students display a good range of academic vocabulary, curriculum programs should incorporate a tangible and substantial vocabulary teaching component.

Keywords: academic vocabulary, vocabulary range, written production, wordlists

Resumen

El presente artículo examina la cantidad de vocabulario académico utilizado por estudiantes de cuarto año en programas de Bachillerato en Inglés y Bachillerato en la Enseñanza del Inglés. El autor explica cómo el vocabulario académico puede ser medido y los beneficios que esto conlleva. El artículo contiene un breve análisis de los resultados del vocabulario académico empleado en 46 trabajos finales de investigación. Basado en estos resultados, el autor aclara que, aunque los estudiantes demuestran un buen rango de vocabulario académico, el currículo debe incorporar componentes tangibles y sustanciales para la enseñanza de dicho vocabulario.

Palabras clave: vocabulario académico, rango de vocabulario, producción escrita, listas de palabras

Introduction

The role of vocabulary has been essential from the very beginning of second language teaching and learning. However, vocabulary instruction gained momentum with the advent of computing resources and the creation of digital corpora around the 1970's (Nation, 2016). Since then, compiling word lists from different fields, as well as using those lists to teach second language learners, has been a breakthrough in course programs, curriculum design, and material development adjusted to better suit students' needs. One of such lists is the Academic Word List developed by Coxhead in 2000. According to Nation (2016), the Academic Word List was intended to be used "with students about to begin university study in an English-speaking country with the primary purpose of reading academic texts" (p. 11). The main assumption here, Nation explains, was that students at that level were already familiar with more basic lists such as the General Service List.

However, it is not just reading academic texts that should be the main focus of the Academic Word List. In Costa Rica, a country whose native language is Spanish, university students are expected not only to be somewhat proficient in one or two linguistic skills but to interact with the language at a professional level. In particular, students who study English or English teaching need a solid vocabulary foundation. Nadarajan (2007, p. 90) mentions that "academic vocabulary [should be] the learning goal of the adult L2 language classroom, [since it] is common across a wide range of academic texts,

but it is not so common in non-academic texts." In addition, Sutarsyah, Nation, and Kennedy (1994) found that academic vocabulary comprises 8.4% of words in the Lancaster-Oslo/Bergen Corpus (LOB) and Wellington corpora and 8.7% of words in economics texts. Coxhead (2006) also states that academic vocabulary instruction is important because

- […] understanding and properly using this vocabulary allows students to be part of the academic community.
- To be successful at university, learners need to be able to show that they can read, understand and respond clearly in writing and speaking to academic language and concepts.
- Vocabulary is a very important part of literacy.
- Students will meet general academic words many times in their academic reading.
- These words occur in a wide variety of subject areas (p. 3).

For these reasons, vocabulary lists should be introduced into the curriculum to strengthen students' linguistic competence. In this way, students will benefit by getting both planned and informed direct and indirect exposure to the vocabulary necessary to develop professionally and conduct research in different fields.

Therefore, the purpose of this study is to explore the amount of academic vocabulary present in students' final research paper. Currently, vocabulary lists and direct vocabulary instruction are not explicitly implemented in the BA in English or BA in English Teaching. Because of this, vocabulary

recycling is not currently addressed, and this is an essential component in mastering vocabulary (Coxhead, 2006). Coxhead (2006) also mentions that language programs and courses should prioritize vocabulary learning and teaching. Vocabulary acquisition should, therefore, should have clear and attainable goals. Understanding students' use of academic vocabulary will shed light on whether explicit vocabulary lists should be part of the language program, whether they should be included more explicitly, or whether they play no major role in students' writing production. Analyzing this data will benefit students enrolled in the English and English Teaching majors since the results will have a direct impact in the curriculum.

Literature Review

Word features. Authors such as Nation (2001), Coxhead (2006), and Akmajian *et al.* (2010) suggest that, to determine vocabulary acquisition, students should know different features of a word. First, students should be familiar with form. Word form deals with pronunciation, spelling, and word parts (e. g. understanding the mechanics of roots, prefixes, and suffixes). Second, students should understand the meaning of a word. This comprises knowing how the word can be defined and being able to identify several meanings of the word in various specific linguistic contexts. Finally, it is necessary to understand the use of a word. This refers to word collocations, use constraints, and grammatical functions. Therefore, words should meet the different criteria established above to be considered as correct. Gains and Redman (2000)

also propose a difference between *receptive* and *productive* vocabulary. Receptive vocabulary is that in which "language items can only be recognized and comprehended in the context of reading and listening material, and productive vocabulary [is that] which the learner can recall and use appropriately in speech and writing" (Gains and Redman, 2000, p. 64). Subsequently, appropriate instruments need to be established to know if students have mastered a given set of words in the receptive or productive domain. Several authors (Nation, 2001; Coxhead, 2006; Daller, Milton, and Treffers-Daller, 2007; Read, 2004) believe that pen and paper tests are a means to know if students have acquired vocabulary. They also mention paragraph or essay writing to measure students' word knowledge. From the different assessment instruments proposed, this investigation favors extensive academic writing since the focus is academic vocabulary and since it stands as the best procedure to determine if students have acquired "productive" vocabulary and the components of form, meaning, and use mentioned above.

Vocabulary measurement. In order to measure vocabulary, it is necessary to decide what and how words would be counted. To achieve this, a clear distinction between word categories should be made. Nation (2001) mentions four possible pathways to count words.

1. Tokens: The total number of tokens corresponds to all the words used in a text. Every word counts as a token. This includes words that are repeated in the same text, be it oral or written. For example, a sentence

like "I do not know why I called you." would contain eight words or tokens. Even though the word "I" is repeated, it counts as two separate tokens.

2. Types: In this case, words that are repeated are counted as just one instance. In the sentence "I do not know why I called you." we have seven types since the word "I" is repeated, and it counts as one type only.

3. Lemmas: A lemma consists of any headword plus some inflected or reduced forms (e. g. do not and don't). Because it deals with inflectional suffixes, lemmas belong to the same part of speech (e.g. write, writes, writing, written).

4. Word families: A word family includes the headword, inflected forms, and derived forms. For example, the common expression "We agreed to disagree." Contains only three word families since "agreed" and "disagree" belong to the same family.

Because lemmas are somewhat limited and word families already include them, the analysis conducted in this paper will include tokens, types, and word families only.

Measuring vocabulary density has often been intended to create language corpora or analyze children's vocabulary acquisition. The implications in the field of applied linguistics have undoubtedly been considerable, and they range from the development of new methodologies to the creation of textbooks and materials that include different kinds of vocabulary. Nevertheless, research has been usually conducted on how to teach vocabulary from corpora or the evaluation of vocabulary in various contexts. Except for developmental vocabulary acquisition in children, vocabulary size, as such, has not been widely studied in adults from an academic, writing perspective.

In terms of English vocabulary acquisition, other scholars have also focused on the breadth of vocabulary knowledge (how many words learners know) and depth of vocabulary knowledge (how well learners know words) in particular groups of students. One study of this kind was carried out by Santos (2003). Using a sub-sample of ten students from a total of 104 college students enrolled in an urban New England community college, the author sought to measure students' breadth and depth of vocabulary in English. All students in the study came from language-minority backgrounds, such as Spanish, German, and Portuguese, among others. To measure breadth and depth of vocabulary knowledge, various test items and sentence production exercises were used. As part of her findings, Santos (2003) mentions the relationship between students' L1 academic proficiency and students' academic vocabulary knowledge in English. On the other hand, Santos states that when students showed a greater breadth of academic vocabulary knowledge, they were able to identify more multiple meanings for a single word in English. Finally, the author mentions how students' linguistic background influences students' performance. Students whose native language was Spanish had an advantage over native speakers of Chinese, for example. The reason for this is the number of cognate words that translate into high-frequency words in one language but not in others.

Similarly to Santos (2003), Jong (2008) conducted some case-study

research to find out the level of mastery developed by Korean students learning academic vocabulary in English. Particularly, the study focuses on "accuracy and control of the words they use in their English compositions." (p. 4). Three Korean students enrolled in a composition course were selected based on the length of their stay in the U.S. and the duration of their English studies. For two semesters, these students submitted eight papers from their English composition classes plus two papers they wrote for other classes on different topics. The analysis was done through the Vocabulary Knowledge Scale (VKS), vocabulary interviews, and retrospective protocol interviews. All subjects improved, in various degrees, both vocabulary recognition and use. In addition, Jong (2008) concluded that four factors stood out as more problematic for the students' English vocabulary learning: "a) abstractness and multiplicity in the meanings of English academic words; b) difficulty in retrieving words and shortage of necessary vocabulary; c) difficulty in using the bilingual dictionary as a reference [; and] d) learners' view on the role of vocabulary learning and their personal desire for academic success" (p. 234). Because the primary purpose of the present research is measuring students' academic vocabulary, how students process learning will not be analyzed here.

As in the previous study, other researchers have studied the connection between vocabulary and proficiency. Olmos (2009, p. 73), for example, points out that "several research studies have proven a direct relationship between the number of words known by a foreign language student and his/her language proficiency." In her study, Olmos (2009) sought to determine the correlation between the vocabulary students have studied and the vocabulary they have retained. In order to come up with this, she assessed students who speak Spanish as their first language using Vocabulary Levels Tests, which are a series of matching tests where subjects must select the correct definition or synonym for three words from one of six options (Webb, 2018). After administering the tests, students' scores were analyzed and different correlations were established to see if students' grades were consistent with their level. Out of the 49 students who took the basic tests, only three students reached the minimum score required to pass. In the case of the 38 students who took the more advanced tests, none of the students reached the minimum score. According to the author, the results of the study indicate that vocabulary retention "in Spanish high schools does not reach the minimum standards established to make these students efficient language users" (p. 87). To improve this situation, Olmos (2009) suggests that certain changes must be made in the manner input is presented to students, paying particular attention to frequency levels, students' age, and academic stage. Finally, she stresses that the curriculum should incorporate different vocabulary teaching techniques to improve students' vocabulary acquisition.

Vocabulary production. Besides measuring the amount of vocabulary students use, it is important to determine how to improve students' vocabulary production. In this sense, two main approaches to word teaching and

learning have been suggested: incidental and intentional. Atzler (2011, p. 22) comments that "the distinction between incidental and intentional corresponds to the implicit-explicit […domain, where] explicit vocabulary teaching approach is inevitable, especially in the early stages of students' vocabulary learning." According to Read (2004, p. 147), the central issue is "the extent to which learners can acquire word knowledge […as] a by-product of their main learning activity inside or outside the classroom, rather than through activity that is primarily intended to enhance their vocabulary knowledge." The author adds that the communicative approach geared toward that access to sufficient comprehensible input would lead to automatic vocabulary acquisition; nevertheless, more recent research has widely refuted this position. On the other hand, other experts have suggested that it "is the quality and frequency of the information processing activities (i.e., elaboration on aspects of a word's form and meaning, plus rehearsal) that determine retention of new information" (Hulstijn, 2001, p. 275). Clearly, despite the approach preferred, a need for an informed procedure exists and explicit vocabulary teaching becomes indispensable. The researcher believes that language curricula should be clear regarding what vocabulary to teach and how to teach it, and that this will affect students' linguistic performance.

Other researchers, such as Chung (2003), Hirsh and Nation (1992), Klinmanee and Sopprasong (1997), and Laufer and Nation (1995) have explored vocabulary production using vocabulary analysis programs such as Range and Frequency. In particular, Laufer and Nation (1995) sought to establish "the reliability and the validity of the Lexical Frequency Profile as a measure of lexical richness in free written production" (p. 313). This study included 65 participants from different backgrounds and native languages (Chinese, Japanese, Thai, Samoan, Polish, Malay, Russian, and Hebrew). To gather the necessary data, students were asked to write two 300-350-word, in-class compositions. Topics for this task dealt with general matters and controversial issues. The experts then introduced all compositions into the software called VocabProfile (now Lextutor). Among the different results found, Laufer and Nation (1995) concluded that "the Lexical Frequency Profile has been shown to be a reliable and valid measure of lexical use in writing" and that "we can reasonably expect learners' vocabulary size as measured by a vocabulary test to be reflected in the learners' productive use of the language" (p. 319). This evidence suggests that although using vocabulary tests remains a valid choice for measuring vocabulary range, productive skills also stand as a paramount option in lexical analysis.

Method

Participants. A personal electronic mailing list of 46 students taking the last writing course from the B.A. in English and B.A. in English Teaching was created in both the first and second semesters of the academic year. The list consisted of students who were taking the course for the first time and agreed to participate in the study. All students were sent emails asking them to submit their final papers.

A total of 33 papers were received. Data from these papers were collected and analyzed. No paper was kept from analysis. In case students had participated before and were taking the course for the second time, they were explicitly asked not to send their final paper again.

The population included in this study belongs to an undergraduate, ESL program from a public university. Different from other language courses, students do not only learn the language but the rules that govern it. Therefore, at the end of the program, it is expected that they are able to use the language proficiently, and it is also expected that they can describe the language from a linguistic point of view. By the end of the major, students are required to write an academic research paper. All students share Spanish as their native language.

Materials. A written consent document (Appendix 1) was created and distributed among the students encouraging them to participate. An electronic mail was also sent as a means to remind students to send their final essay. Students' final essays were stored electronically and were slightly modified to delete words that were not part of the study (e.g. students' names, institution, and direct quotes). The researcher did not modify participants' papers in any other way. To analyze data, the software package called Range and Frequency was used. This software was programmed by Alex Heatley and designed by Paul Nation and Averil Coxhead of the School of Linguistics and Applied Language Studies at Victoria University. When running *frequency* mode, "the output is an alphabetical list, or a frequency ordered list. It gives the rank order of the words, their raw frequency and the cumulative percentage frequency" (p. 4). In *range* mode, users can "compare a text against vocabulary lists to see what words in the text are and are not in the lists, and to see what percentage of the items in the text are covered by the lists" (p. 2). For the purposes of this study, only *range* mode will be used. The researcher would like to acknowledge that the software can be adapted, distributed, and used freely.

Procedure. This study used a quantitative study design. The researcher asked for permission to attend a class and ask students for their participation. After briefly explaining the nature of the study, students were given a written consent. The researcher instructed them to read it, ask any questions they considered necessary, and sign it if they wished to participate. After collecting all the consents, a list of 46 participants was created. The first electronic mailing was sent to all 46 participants. The electronic mail included a copy of the written consent form addressed to the participants. In this email, the researcher asked students to send their final paper when they had it ready. A second electronic mail was sent approximately one week after the first one. The purpose of this second email was to thank those participants who had already sent their final papers and to encourage the ones who had not submitted it to do it promptly. After two weeks, a final mail was sent thanking all students for their participation and offering them research results. After collecting all papers, the researcher deleted all unnecessary information like proper names and direct

quotes since it was obvious that these particular segments of the text contained irrelevant data or had not been written by participants themselves. In cases of paraphrasing, the text was left untouched since the nature of paraphrasing is to report information using one's own words.

Analysis of the Results

To explore the amount of academic vocabulary present in students' final research papers, the researcher determined the frequency of word use through RANGE and FREQUENCY software. As can be seen from Table 1, with the exception of list 3 and in a lesser degree list 5, students' academic vocabulary use tends to decrease if only tokens are taken into account. Considering that the sublists are based on the frequency of occurrence of the words in the Academic Corpus, this tendency remains between regular parameters. On the other hand, the total amount of academic vocabulary used in terms of tokens accounts for nearly 10.89%. According to Coxhead (2000, p. 213), "the Academic Word List (AWL) contains 570 word families that account for approximately 10.0% of the total words (tokens) in academic texts." Thus, students writing shows great similarities to the academic corpus developed by Coxhead in terms of size and tokens per word list. In terms of types, the tendency displays similar parameters to those of tokens, but, in this case, a sudden increase occurs in sublist 6. This means that the repetition of words from list 6 was comparatively greater than that in any other list with more tokens.

Table 1

Text coverage based on the ten AWL subclasses

WORD LIST	TOKENS/%	TYPES/%	FAMILIES
one	6046 / 3.47	234 / 3.58	57
two	4429 / 2.54	186 / 2.85	58
three	1456 / 0.83	158 / 2.42	55
four	1945 / 1.12	149 / 2.28	56
five	1559 / 0.89	127 / 1.94	50
six	1694 / 0.97	134 / 2.05	49
seven	691 / 0.40	106 / 1.62	50

eight	699 / 0.40	105 / 1.61	47
nine	292 / 0.17	73 / 1.12	42
ten	170 / 0.10	34 / 0.52	17
not in the lists	155438 / 89.12	5228 / 80.01	?????
Total	174419	6534	481

Source. Self elaboration using Range and Frequency software. The software automatically includes question marks (?????) when data is impossible to analyze since they are not part of the Academic Word List.

To extrapolate these results, the researcher also examined data from each of the sublists to determine how many word families were not used. As shown in Table 2, excluding list 10, which consists of thirty word families only, all sublists are composed of sixty word families. A total of 89 (15.63%) word families were not used at least once in students' essays. From this data, it is possible to conclude that three major divisions of use exist. In sublists 1, 2, 3, and 4, students used more than 94.10% of word families on average, but the percentage of word families employed drops to 82.77% in sublists 5, 6, and 7 and to 70.66% in sublists 8, 9, and 10. This finding indicates that there is a relevant segment of word families that is being omitted, which in turn may limit students' writing fluency, accuracy, and complexity.

Table 2

Word Family Use Ratio

WORD LIST	TOTAL WORD FAMILIES	FAMILIES USED	PERCENTAGE OF FAMILIES USED
one	60	57	95%
two	60	58	96.66%
three	60	55	91.66%
four	60	56	93.33%
five	60	50	83.33%
six	60	49	81.66%

seven	60	50	83.33%
eight	60	47	78.33%
nine	60	42	70%
ten	30	17	56.66%
Total	570	481	84.38%

Source. Self elaboration using Range and Frequency software.

It is important to mention that Range and Frequency software also displays all the word families of each sublist, the word families students used, the frequency of the word type, and the frequency of the word family. Since the main purpose of this study was to measure range, this information has been purposely omitted. Nevertheless, in appendix 2, the researcher lists all the word families from the AWL corpus that students did not use.

Conclusions

Academic writing has become a very important component of university life and career success. Many majors require academic writing and more jobs demand fluent and formal writing skills. The trending interdisciplinary nature of the job market requires written communication to be efficient and precise. As Gains and Redman (2000) put it, "students who are required to read technical reports in their native country will have different lexical needs than those who want survival English for travel purposes" (p. 59). Therefore, language academic programs should develop clear policies in terms of vocabulary acquisition.

A first step in this direction should address both students' needs and current vocabulary level.

As part of vocabulary level analysis, the results of this study provide important information about students' vocabulary use in formal academic writing. First, a positive correlation exists between what students wrote and what corpora demonstrates is produced in out-of-class, academic contexts. As a whole, students displayed a good range of academic vocabulary in their final papers. Most of the data per sublist resembles what is normally found in other academic papers. Moreover, another similarity occurs when considering the entire lexicon present in the analysis. A balance between academic and non-academic or general vocabulary resembles the word distribution found by Coxhead (2000) using the academic corpus.

In contrast, from sublist 3 to 10, word use ranks higher in types but low in tokens. The vocabulary from wordlist 3 in particular displays a high use of types and families but ranks very low in terms of word tokens. According to Thomas (2005), this is the result of using the same word types over and over again; consequently, the text lacks lexical richness. The researcher must acknowledge that this discrepancy

between types and tokens is significant in sublist 3 only; therefore, it must not be seen as a major issue but as a call for instructional awareness.

Another aspect that deserves attention is word family use ratio. On one hand, 89 word families (see appendix 2) do not appear in any of the research papers. On the other hand, word family frequency decreases somewhat drastically from sublist 1 to sublist 10. This poses certain questions which need to be answered in order to develop a sound academic vocabulary teaching program: Are these word families difficult due to students' level? Do students consider these word families unimportant or too fancy? Were students ever exposed to these word families in some way or do they ignore that they exist or do not know how to use them? By knowing that a gap exists, instructors can develop adequate policies to determine which parts of the curriculum deserve more attention.

A proposal of this nature should include vocabulary analysis of students' writing and oral skills, principles of learning vocabulary, teaching academic vocabulary, and indirect and direct learning strategies, among others. The instruction as a whole should develop and communicate clear guidelines for vocabulary learning, which includes academic vocabulary. Professors should be encouraged to develop materials and activities aimed at improving students' receptive vocabulary, "that which can only be recognized and comprehended in the context of listening and reading material", and productive vocabulary, "that which the learner can recall and use appropriately in speech and writing" (Gairns and Redman, 2000, p. 64). Finally, the curriculum should lead

students to become autonomous vocabulary learners. They should develop a culture of vocabulary inquiry and awareness, giving vocabulary instruction the same importance as grammar, writing, or pronunciation.

In order to establish a solid curriculum that formally incorporates vocabulary acquisition, further research is required to determine the general needs of the population and the frequency and use of academic words. Research about the amount of academic vocabulary in oral production should also be taken into account. This will allow comparing and contrasting its use and standardize pedagogical procedures to increase students' vocabulary across skills, taking into account incidental and explicit repetition. Finally, research should also seek to ascertain the strengths and limitations of the new curricular adaptations and pedagogical practices.

Bibliography

Akmajian, Adrian; Demers, Richard; Farmer, Ann; Harnish, Robert. (2010). *Linguistics: An Introduction to Language and Communication* (6th edition). United States of America: The MIT Press.

Atzler, Judith Kerstin. (2011). T*wist in the list: Frame Semantics as a vocabulary teaching and learning tool. Dissertation (Doctor of Germanic Studies).* Austin, TX: The University of Texas at Austin.

Chung, Teresa. (2003). A corpus-comparison approach for term extraction. *Terminology 9*, 2: pp. 221-246.

Coxhead, Averil. (2000). A New Academic Word List. *TESOL Quarterly, 34*(2): pp. 213-238.

Coxhead, Averil. (2006). *Essentials of Teaching Academic Vocabulary*. Boston: Houghton Mifflin.

Daller, Helmut; Milton, James; & Treffers-Daller, Jeanine. (2007). *Modelling and Assessing Vocabulary Knowledge*. Cambridge: Cambridge University Press.

Gains, Ruth & Redman, Stuart. (2000). *Working with Words*. Cambridge: Cambridge University Press.

Heatley, Alex; Nation, Paul; & Coxhead, Averil. (2002). Range and Frequency programs. http://www.vuw.ac.nz/lals/staff/Paul_Nation

Hirsh, David; & Nation, Paul. (1992). *What vocabulary size is needed to read unsimplified texts for pleasure? Reading in a Foreign Language 8,* 2: pp. 689-696.

Hulstijn, Jan (2001). Intentional and incidental second language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In P.

Robinson (Ed.), *Cognition and second language instruction* (pp. 258–286).Cambridge: Cambridge University Press.

Jong, Youngkyong. (2008). Learning and use of academic vocabulary: A case study of three korean undergraduate students in a U.S. academic setting (Order N.° 3326345). Available from ProQuest Central; ProQuest Dissertations & Theses Global. (304473542). Retrieved from http://search.proquest.com.ezproxy.sibdi.ucr.ac.cr:2048/docview/304473542?accountid=28692

Klinmanee, Nanta; & Sopprasong, Lawan. (1997). Bridging the vocabulary gap between secondary school and university: a Thai case study. *Guidelines 19,* 1: pp. 1-10.

Laufer, Batia; & Nation, Paul. (1995). Vocabulary Size and Use: Lexical Richness in L2 Written Production. *Applied Linguistics, 16*(3), p. 307. https://www.learntechlib.org/p/79856/

Nadarajan, Shanthi. (2007). Measuring academic vocabulary size and depth in the writing classroom: Does it really matter? (Order N.° 3284352). Available from ProQuest Dissertations & Theses Global. (304894148). Retrieved from http://search.proquest.com.ezproxy.sibdi.ucr.ac.cr:2048/docview/304894148?accountid=28692

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

Nation, Paul. (2016). *Making and using word lists for language learning and testing*. Amsterdam: John Benjamins Publishing Company.

Olinghouse, Natali;, & Wilson, Joshua. (2013). The relationship between vocabulary and writing quality in three genres. *Reading and Writing, 26*(1), pp. 45-65. doi:http://dx.doi.org/10.1007/s11145-012-9392-5

Olmos, Carmen. (2009). An assessment of the vocabulary knowledge of students in the final year of secondary education. Is their vocabulary extensive enough? *International Journal of English Studies,* pp. 73-90. Retrieved from https://search-proquest-com.ezproxy.sibdi.ucr.ac.cr/docview/808151700?accountid=28692

Read, John. (2004). 7. Research in teaching vocabulary. *Annual Review of Applied Linguistics, 24,* pp. 146-161. Retrieved from https://search-proquest-com.ezproxy.sibdi.ucr.ac.cr/docview/197926905?accountid=28692

Santos, Maricel. (2003). Three studies on the academic vocabulary knowledge of United States language-minority community college students (Order N.° 3100161). Available from ProQuest Dissertations & Theses Global: Literature & Language; ProQuest Dissertations & Theses Global: Social Sciences. (305335610). Retrieved from http://search.proquest.com.ez-proxy.sibdi.ucr.ac.cr:2048/docview/305335610?accountid=28692

Sutarsyah, Cucu; Nation, Paul; & Kennedy, Graeme. (1994). How useful is EAP vocabulary for ESP? A corpus based case study. *RELC Journal, 25*(2), pp. 34-50.

Thomas, Dax. (2005). Type-token Ratio in One Teacher's Classroom Talk: An Investigation of Lexical Complexity. United Kingdom: University of Birmingham.

Webb, Stuart. (2008). Receptive and productive vocabulary sizes of l2 learners. S*tudies in Second Language Acquisition, 30*(1), pp. 79–95. http://doi.org/10.1017/S0272263108080042

Appendix 1

UNIVERSITY OF COSTA RICA INFORMED CONSENT STATEMENT FORM

Fourth Year Students' Use of Academic Vocabulary

You are invited to participate in a research study about the use of academic vocabulary by fourth year students at Lenguas Modernas. You were selected as a possible subject because you are taking the last writing course from the English, core block at the University of Costa Rica. I ask that you read this form and ask any questions you may have before agreeing to participate in the study.

This study is being conducted by William Charpentier Jiménez, English professor at the University of Costa Rica in the School of Modern Languages, Rodrigo Facio Branch, San José, Costa Rica.

Purpose of the study:

The purpose of this study is to help us understand current practices in vocabulary acquisition as well as vocabulary use by students. This information will shed light on the strengths and weakness in students' preparation during the major in terms of vocabulary acquisition and writing skills. The study will focus on students who are taking Rhetoric IV during 2015.

Number of people taking part in the study:

All B. A. students in English and English Teaching from the University of Costa Rica who take Rhetoric IV in 2015. The exact number of people who will choose to participate in the study is unknown.

Procedures for the study:

For the study, you may participate, without any obligation to do so, by sending a copy of your final paper to wcharpentier@gmail.com

Participating in the study will require very little time. You do not need to modify your final paper, except for deleting your name from it and sending it through email. Doing this may require approximately 10 minutes. This time will be scheduled at your convenience. You have up to two weeks to send your paper after you receive the invitation email. You will get one notification and two reminders.

Risks of taking part in the study:

The risks of participating in this study are nominal. The only foreseeable risk is the chance of being identified, due to the nature of the delivery mechanism of information. Measures will be taken however to ensure confidentiality. If at any time, you feel uncomfortable or do not want to participate for any reason, you can tell the researcher.

Benefits of taking part in the study:

There is no direct benefit that can be reasonably expected by your participation in this research. The benefits of this study are indirect and include what may be learned about the use of vocabulary by students who have advanced studies in the English language.

Confidentiality:

Efforts will be made to keep your personal information confidential. The researcher cannot guarantee absolute confidentiality. Your personal information may be disclosed if required by law. Your identity will be held in confidence in anything the researcher writes, publishes, presents, or shares with others about the study. All of the data will be stored securely for an indefinite period of time and only the researcher will have access to these materials. If necessary, the researcher will use pseudonyms (names he makes up) for all the participants in the study and any other identifying information.

Payment:

You will not receive payment for taking part in this study.
Contacts for questions or problems:

For questions about the study, contact the researcher, William Charpentier, at 88489182 in Costa Rica, or by e-mail at wcharpentier@gmail.com.

Voluntary nature of the study:

Taking part in this study is voluntary. You may choose not to take part or may leave the study at any time. Leaving the study will not result in any penalty or loss of benefits to which you are entitled. If you withdraw from the study before data collection is completed, all the information collected during the study about you (as discussed in the Procedures for the Study section) will be destroyed. Your decision whether or not to participate in this study will not affect your current or future relations with the University of Costa Rica.

Consent:

In consideration of all of the above, I give my consent to participate in this research study. I will be given an electronic copy of this informed consent document to keep for my records. I agree to take part in this study and to any further analysis of my written work.

Subject's Name: _____

Subject's Email: _____

Subject's Signature: _____ Date: _____

Printed Name of Person Obtaining Consent: <u>William Charpentier Jiménez</u>

Signature of Person Obtaining Consent: _____ Date: _____

Appendix 2

MISSING WORD FAMILIES IN ALL PAPERS

Sublist 1:
Export, income, legislate
Sublist 2:
Credit, equate
Sublist 3:
Constrain, emphatic, fund, layer, remove
Sublist 4:
Civil, domestic, parameter, regime
Sublist 5:

Amend, clause, compound, discrete, liberal, license, ratio, revenue, substitute, welfare

Sublist 6:

Aggregate, allocate, estate, federal, fee, interval, migrate, overseas, rational, subsidy, underlie

Sublist 7:

Channel, chemical, empirical, extract, finite, mode, paradigm, prohibit, reverse, submit

Sublist 8:

Arbitrary, commodity, currency, deviate, displace, inspect, offset, plus, practitioner, radical, thereby, uniform, widespread

Sublist 9:

Analogy, behalf, bulk, cease, commence, concurrent, erode, ethic, medium, military, overlap, preliminary, revolution, rigid, route, subordinate, suspend, violate

Sublist 10:

Adjacent, albeit, assemble, collapse, conceive, integrity, invoke, levy, odd, ongoing, so-called, straightforward, whereby