

7th International Conference on Corpus Linguistics: Current Work in Corpus Linguistics:
Working with Traditionally-conceived Corpora and Beyond (CILC 2015)

Analysing lexical density and lexical diversity in university students' written discourse

Carmen Gregori-Signes^a, Begoña Clavel-Arroitia^{b*}

^aUniversitat de València-IULMA, Blasco Ibáñez 32, Valencia 46010, Spain

^bUniversitat de València-IULMA, Blasco Ibáñez 32, Valencia 46010, Spain

Abstract

This study analyses both lexical density and lexical diversity in the written production of two groups of first year students at the Universitat de València at the beginning and end of one-semester teaching period. These results were compared with those obtained by a third group of students aiming at level C2. Lexical density was tested using *Textalyser* (<http://textalyser.net>) and lexical frequency used the software RANGE (Nation and Heatly, 1994). Our results prove that the students from both groups at level B1 show the same progression between writing tasks 1 and 3. Furthermore, we can claim that it is possible to obtain a reliable measure of lexical richness which is stable across two pieces of writing produced by the same learners.

© 2015 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Peer-review under responsibility of Universidad de Valladolid, Facultad de Comercio.

Keywords: Lexical density; Lexical diversity; Learner corpus; English Language Teaching.

1. Introduction

Lexical density, lexical diversity or lexical richness (Daller, van Hout & Treffers-Daller, 2003) are terms which refer to statistical measures that gauge the lexical richness of texts and may also be used to assess students' overall progress. The lexical richness of a text accounts for how many different words are used in a text, while lexical

* Corresponding author. Tel.: +0-34-96-3864262; fax: +0-34-96-3864161.

E-mail address: begona.clavel@uv.es

density provides a measure of the proportion of lexical items (i.e. nouns, verbs, adjectives and some adverbs) in the text (Johansson, 2008: 61). Both measures have been applied in computer analyses of corpus data. As a rule, texts with a lower density are more easily understood and spoken texts have lower lexical density levels than written texts (Ure, 1971; Halliday, 1985). However, as argued by Johansson (2008), a text may have high lexical diversity (i.e. contain many different word types), but low lexical density (i.e. contain many pronouns and auxiliaries rather than nouns and lexical verbs) or vice versa.

This article focuses on both lexical density and lexical diversity, by analyzing the Lexical Frequency Profile (Laufer and Nation, 1995) in two corpora of four essays written by two different groups of students aiming at level B1, and a third one by students aiming at level C2 used for comparison purposes. The objective is twofold. On the one hand, to test the validity of the programme RANGE for the detection of lexical progression and, on the other, to determine whether a short span of six weeks is sufficient to start detecting the progress of a group of students.

The corpus consists of two pieces of writing by each B1-student, which amounts to a total of 108 essays ranging from 150 to 250 words, and 27 essays by C2 students. The students were officially registered in a Preliminary English Test (PET) course part of the degree in English Studies at the Universitat de València. Accordingly, the writing tasks were designed to fit the requirements of the PET (B1 in the CEFR) official examinations. The first assignment (Corpus A) was collected during the second week of class (writing task 1); and the second (Corpus B) six weeks later (writing task 3), after the students had been exposed to regular teaching in English. Their Lexical Frequency Profile (LFP) was compared to one sample of writing produced by C2 students (Corpus C) also enrolled at the University of Valencia. The objective of this study was to find out the degree of lexical richness at the beginning and the end of the teaching period. That is, a) if there were any (significant) differences after a four-hours per week instruction and exposure to the language in other content subjects, over a six-week period and; b) if there were any significant differences between corpus A, corpus B and corpus C.

The results obtained have been incorporated into a database compiled by the group CASTLE[†] (Compilación del Corpus de Aprendices de Inglés como segunda y tercera lengua de la Universidad de Valencia). CASTLE is a research project whose objective is to collect a longitudinal corpus designed to study lexical development in the same group of students over a period of four years, in order to find out possible patterns in their progression. As Hawkins and Filipovic (2012: 134) highlight:

“documenting the first appearance in a corpus is straightforward. Analysing and classifying the semantic progression is much less so. Indeed, any attempt to do a serious semantic analysis of the expanding lexicon of learners raises all the deep issues that are currently being researched [...] in the various branches of lexical semantics and lexicography”.

The programme RANGE was used to obtain the LFP in the three corpora. RANGE provides a range or distribution figure (how many texts a particular word occurs in), a headword frequency figure (the total number of times the actual headword type appears in all the texts), a family frequency figure (the total number of times the word and its family members occur in all the texts), and a frequency figure for each of the texts the word occurs in.

As reported by studies like that by Goodfellow, Lamy and Jones (2002), the LFP cannot be fully trusted to assess the “learners' vocabulary knowledge in the same way that Laufer & Nation's study claimed to do”. However, the same authors admit that its results “correlate significantly with scores awarded by human markers for vocabulary use” in the texts they analysed and suggest that the approach could be used for self-assessment, among other possible applications. A similar principle applies here since ours is one of the many preliminary tests which are

[†] (CASTLE-UV-INV-PRECOMP12-80754; CASTLE-GV/014/022).

being carried out with data obtained for the project CASTLE which intends to measure our students' progress and find reliable tools to measure it. Among its objectives, the project aims to determine the shortest span of time needed to detect differences in students' written production and indicate if progress has been made. Since teachers have observed that most students do in fact improve during the first three months of their degree, the objective was to test if this progress was real and if it showed when using statistical tools to measure it, thus confirming the teachers' intuition about their progress. Pedagogical reasons support this decision, since, as has often been argued, students at higher levels do not depend exclusively on classroom input but learn from other sources. Thus, a three-month period in which the students are exposed to teaching in English may well be sufficient to show some learner progress. As claimed by Laufer and Nation's (1995), two essays by the same learner are enough to assess lexical richness if we measure it through *the Lexical Frequency Profile*, "which looks at the proportion of high frequency general service and academic words in learners' writing" (Laufer & Nation, 1995: 307).

Students' written production has always been a central part in the assessment of their linguistic competence, therefore knowledge of lexical richness obtained through reliable quantitative and qualitative measures can provide teachers with a more accurate picture of lexical progress, at least in the case of whole groups (Webb & Paribakht, 2015) and may help teachers to reflect on their teaching and the suitability of their teaching materials. As for students, lexical richness, as concluded by Goodfellow, Lamy and Jones (2015) "could be a first step in the development of a viable aid for learners evaluating their own writing."

2. Lexical diversity/richness

Laufer and Nation (1995: 309) list several measures for measuring lexical richness: lexical originality (LO), lexical density (LD), lexical sophistication (LS) and lexical variation (LV). The LO index "measures the learner's performance relative to the group in which the composition was written." Laufer and Nation (1995) argue that such a measure is not reliable since it is defined not only by the composition in question but by the group factor (LO is calculated by multiplying the number of tokens unique to one writerx100 and dividing it by the total number of tokens); b) LD is defined as the percentage of lexical words in the text, i.e., nouns, verbs, adjectives, adverbs; c) LS is the percentage of 'advanced' words in the text (N° of advanced wordsx100/ total number of lexical tokens), while d) LV is the type/token ratio, i.e., the ratio measured as a percentage between the different words in the text and the total number of running words ($\text{number of words} \times 100 / \text{number of tokens}$). LV has proved to be unstable for short texts and can be affected by differences in length. This article focuses on LD and LV.

2.1. Lexical density

Measuring vocabulary development has been, and still is, a major object of research in the field of EFL. As reported by McKee et al. (2000), many commonly used measures have been based on the ratio of different words (Types) to the total number of words (Tokens), known as the Type-Token Ratio (TTR). TTRs can be measured by a large number of computer programmes dedicated to linguistic analysis, such as Systematic Analysis of Language Transcripts (SALT) (Miller & Iglesias, 2008), the Oxford Concordance Programme (Martin & Hockey, 1988), and the Computerized Language Analysis (CLAN) programmes (MacWhinney & Snow, 1990; MacWhinney, 1995).

Traditionally, lexical diversity has been measured using the TTR. A text is considered 'dense' if it contains many lexical words relative to the total number of words, i.e., lexical and functional. However, this measure may be problematic according to Johansson (2008: 63) who argues that "it is often necessary to re-use several function words in order to produce one (new) lexical word... [so] a longer text usually gives a lower TTR value than a shorter text." Laufer and Nation (1995: 309) also point out that LD does not necessarily measure lexis "since it depends on the syntactic and cohesive properties of the composition. Fewer function words may reflect more

subordinate clauses, participial phrases and ellipsis, all of which are not lexical but structural characteristics of a composition.”

Relying on LD has been reported as inadequate to measure vocabulary development by several authors who claim that TTRs inevitably fall with the increasing size of the token sample and consequently are not indicative of lexical richness. Thus, “any single value of TTR lacks reliability, as it will depend on the length in words of the language sample used (McKee et al, 2000:325) (cf. also Vermeer, 2000; Johansson, 2008; McCarthy & Jarvis, 2007, 2010). In addition, Granger & Wynne’s (1999) study shows that this [using the type/token results to draw conclusions on lexical richness in learner texts would be very unwise because a learner corpus may contain a very high rate of non-standard forms [...] and these forms may significantly boost the type/token ratios” (Granger, 2002:15). Furthermore, Veermer (2000) emphasizes the fact that we should compare texts on similar topics –since topic changes can strongly affect the results of lexical measures– and suggests controlling the tasks so that the types of topics are similar for all the informants. Veermer (2000) advises researchers never to apply TTR or logTTR, but to count the number of types or lemmas and to relate the words in the data to their difficulty, measured by their frequency (or frequency classes). Therefore, a high type/token ratio indicates how well a learner expresses himself with his own vocabulary and not the types of words he knows since LV distinguishes only between different words, not between the quality of the words used in a composition.

2.2. Lexical Frequency Profile

The LFP, developed by Laufer and Nation (1995: 311), categorises the percentage of words a learner uses at different vocabulary frequency levels and according to which frequency band each word belongs to: first 1.000 most-frequent, second 1.000 most-frequent; with the 570 most-frequent ‘academic’ words not in either of the other two lists. They suggest that for less proficient learners the main distinction should be between the 1.000 first more frequent words, the second 1.000 words or the first two thousand (cf. McCarthy, 1999) and any other vocabulary; while for more advanced learners, the profile could look at the total number of word types of the second 1.000 most frequent words, the academic vocabulary and the less frequent words, i.e., words that are not in the first 1.000 most frequent words and not in any of the above two lists.

The programme RANGE is used to compare the vocabulary of up to 32 different texts at the same time. For each word, it provides a range or distribution figure (how many texts the word occurs in), a headword frequency figure (the total number of times the actual headword type appears in all the texts), a family frequency figure (the total number of times the word and its family members occur in all the texts), and a frequency figure for each of the texts the word occurs in. RANGE offers three ready-made base lists: the first (BASEWRD1.txt) includes the most frequent 1.000 words of English. The second (BASEWRD2.txt) includes the second 1.000 most frequent words, and the third (BASEWRD3.txt) includes words not in the first 2000 words but frequent in upper secondary school and university texts from a wide range of subjects. All the lists include the base and derived forms of each word. The first 1.000 words thus consist of around 4.000 forms or types. The sources of these lists are A General Service List of English Words by Michael West (Longman, London 1953) for the first 2.000 words, and The Academic Word List by Coxhead (1998, 2011) containing 570 word families. The first thousand words of A General Service List of English Words are usually those in the list with a frequency higher than 332 occurrences per five million words, plus months, days of the week, numbers, titles (Mr, Mrs, Miss, Ms, Mister), and frequent greetings (Hello, Hi, etc.).

The programme calculates the LFP in terms of word tokens, word types and word families. According to Laufer and Nation (1995: 312), word families are the most revealing calculation as an indication of lexical richness and the use of *low frequency words* is an indicator of richness too. The advantages of LFP over other measures, according to Laufer and Nation (1995:313ff) can be summarized as follows:

- LFP is largely independent of syntax and text cohesiveness (LD is not).

- LFP provides lists of different types of words, defined by frequency level. LS for example only differentiates between frequent and sophisticated words.
- LFP is independent of learning material, reading lists etc., since it is defined in terms of frequency.
- LFP discriminates between subjects who use less or more frequent vocabulary.

Consequently, comparing the different frequency lists obtained using RANGE will give us information about the lexical richness in the written compositions under analysis.

3. Corpus description

3.1. *The CASTLE Project*

CASTLE represents an innovative idea that originated at the Universitat de València and is one of the few longitudinal learner corpus obtained from Spanish Learners of English as a Foreign Language. The project started in 2012 and its main objective is to compile a corpus made of compositions from students from the Degree in English Studies (English Language 1-8, from B1 to C2 level) and keep a track of the students' linguistic evolution.

The process involves a four-step sampling process per academic year. With the consent of the students, samples one and two are collected at the start and end of the first semester; and samples three and four are collected at the start and end of the second semester. Each sample has around 400 words, is submitted via the university's virtual platform in Word/PDF format, and codified individually. The corpus CASTLE had 1.300 texts and around 500,000 tokens at the end of the first edition and has added 488 texts and 192,388 tokens halfway through the second edition (academic year 2014/15).

3.2. *Procedure for data collection*

Corpora A and B (B1 students) are first-year university students officially registered in a Preliminary English Test (PET) course. Accordingly, the writing tasks were designed to fit the requirements of the PET (B1 in the CEFR) official examinations. Although A2 to B1 is the official level set by the university for the first academic year in English Studies, as reported by Díez-Bedmar's (2012) study of university entrance examinations, it seems that most of the students already enter the university at a B1 level (91.33%).

Table 1 below summarises the corpora collected for this study. The first assignment (Corpus A) was collected during the second week of class; the third task (Corpus B) a month and a half later, after the students had been exposed to regular teaching at B1 level. All students in each group usually write either three or four compositions per term and send them to their teachers as compulsory tasks. For research purposes, we collected the first and the last of those compulsory activities. There were two compositions by each student (writing task 1 and writing task 3).

- Writing task 1 was handed in on 11th of October. The students had only attended two weeks of classes. The task read as follows: "Last week you spent the day in another town with a group of friends. Write a letter to an English speaking friend telling him/her: where you went, who you went with, what you did."
- Writing task 3 was handed in on the 24th of November. The students had already covered most of the examination contents for this first part of the course. The task read as follows: "When applying for a course study (e.g. Erasmus stay) at a university or college, students are expected to write a personal statement in addition to completing an application form. You are applying for one of the universities included in our Erasmus programme. Write your own personal statement."

- Writing task 4 was handed in on January the 4th and corresponds to those essays written by fourth year students. It was one of the writing tasks in the textbook used in the course. Students had to write a letter proposing a city as the venue to hold an international conference.

Table 1. Text Corpora used in this article

Corpus A (C)	27 essays (first task) by Group C
Corpus A (D)	27 essays (first task) by Group D
Corpus B (C)	27 essays (third task) by Group C
Corpus B (D)	27 essays (third task) by Group D
Corpus C	27 essays by C1-C2 students
TOTAL	135 texts

4. Objectives and Methodology

4.1. Aims of the present study

The present study aims at testing the reliability of the software RANGE and the validity of the LFP to measure progress in two compositions by the same students during the same semester. The results obtained are to be incorporated into a database which, in the future, will register the performance by the same students at later stages of their learning process. The steps in the analysis can be summarised as follows:

- Obtain the LFP for Corpus A and B, using the programme RANGE. The objective is to measure any significant progress within a few weeks difference. The results will confirm the level of our students and will allow us to establish a point of departure for the longitudinal study.
- Compare both Corpus A and B (108 essays of 150 to 260 words by B1 students) with Corpus C (C2 students). This should give us significant differences if, as claimed by its creators, RANGE can account for differences when comparing groups from different educational backgrounds and/or systems or groups acquiring language outside the classroom. The results obtained will tell us whether RANGE is adequate to detect improvement in lexis between students who supposedly have a different level.
- Obtain results of LD, readability and the use of vocabulary using the tool *Textalyser* and attempt to identify the existence of signs of progress between the first and the third task handed in by first-year students. The assumption underlying the analysis was that we would find signs of progress when comparing the first- and the fourth-year students writing tasks.

In sum, the object of research was twofold, first to determine whether there are grounds for basing an assessment of the quality of a learner's text on the LFP; secondly, to test the validity of such a short span of time as a valid period for detecting differences among students.

4.2. Tools of analysis

The present article uses the software RANGE, developed by Nation and Heatly (1994) for the analysis proposed here. The reason for having selected this software is due to the availability of the software and to the type of results that it provides for the analysis and comparison of corpora such as the ones analysed here, since our samples are

small. McKee et al. (2000) claim that the software has been designed to fit research where this is the case allowing individual lexical diversity to be measured with accuracy.

The second tool employed in our analysis was *Textalyser*. It is a free software utility, offered by the Bernhard Huber Internet Engineering Company, which allows you to find the most frequent phrases and frequencies of words. It also offers the number of words, characters, sentences and syllables. Finally it calculates LD and readability factors.

5. Analysis and results

5.1. Percentage of tokens, types and families

RANGE calculates the number and percentage of tokens, number and percentage of types and number of families. According to Laufer and Nation (1995: 312), the LFP “uses a definition of what should be counted as a word which most closely matches how learners view words”. The general results for each one of the subgroups are shown in the following table:

Table 2. Compared results for tokens, types and word families

	CorpusA(C)	CorpusA(D)	CorpusB(C)	CorpusB(D)	CorpusC
TOKENS					
One	5108/85.06%	5245/84.26%	5471/88.41%	5433/87.43%	7730/78.29%
Two	337/ 5.61%	359/ 5.77%	253/ 4.09%	271/ 4.36%	596/ 6.04%
Three	39/ 0.65%	33/ 0.53%	176/ 2.84%	186/ 2.99%	568/ 5.75%
Not in the lists	521/ 8.68 %	588/ 9.45%	288/ 4.65%	324/ 5.21%	979/ 9.92%
TYPES					
One	614/55.12%	653/54.87%	632/63.39%	655/61.91%	810/50.37%
Two	147/13.20%	154/12.94%	108/10.83%	129/12.19%	199/12.38%
Three	26/ 2.33%	23/ 1.93%	104/10.43%	106/10.02%	168/10.45%
Not in the lists	327/29.35%	360/30.25%	153/15.35%	168/15.88%	431/26.80%
FAMILIES					
one	436	461	417	431	524
two	126	129	89	100	147
three	24	22	83	87	118
total	586	612	589	618	789

As can be seen in table 2 above, the number of tokens and types obtained with regard to lists one and two show similar numbers. However, if we take a look at the third list, some differences can be observed: first year students seem to use more words belonging to list three in their third piece of work: they go from an average of 24 to an average of 102. There is therefore a meaningful evolution between task 1 and task 3. The results also indicate differences between B1 students and C2 students. Students in the latter group include more words (168) of this type in their compositions. In the case of families, which is one of the most revealing measures, a progression is also observable, in particular in list three, where there is first a progression in B1 students (from 23 to 84) and a difference between B1 students and C2 students, who have 118 in words from list three.

In the case of Corpus C we carried out the analysis again using the UWL as in the third list:

Table 3. Results for corpus C contrasting the academic list and the UWL

	WORD LIST	TOKENS/%	TYPES/%	FAMILIES
ACADEMIC LIST	three	568/5.75	168/10.45	118
UWL	three	235/2.38	71/4.42	71

We can observe differences in the outcomes comparing both lists. As might be expected, there are fewer tokens, types and families in the third list when this is the UWL (5.75% vs. 2.38%, 10.45% vs. 4.42%, 118 vs. 71, respectively).

5.2. Frequent words

Regarding the outcomes concerning the use of particular words coming from the different lists that our students used more frequently, we can observe the general results in table 4:

Table 4. Results for more frequent words

Groups	Lists	Words	Range	Frequency	
Corpus C	List 2	bus	23	27	
		hotels	23	26	
		international	18	35	
		entertainment	18	19	
	List 3	conference	25	55	
		transport	23	44	
accommodation		22	37		
Corpus A(C)	List 2	lunch	13	13	
		dinner	12	15	
	List 3	finally	6	6	
Corpus A(D)	List 2	lunch	9	9	
		dinner	9	9	
		hotel	8	16	
	List 3	finally	5	6	
Corpus B(C)	List 2	improve	18	23	
		lot	14	21	
	List 3	culture	8	10	
		furthermore	9	9	
		research	7	7	
	Corpus B(D)	List 2	improve	16	25
lot			11	15	
program			9	14	
sincerely			8	8	
List 3		culture	14	18	
		academic	7	8	
		cultures	6	7	

The vocabulary used by students is undoubtedly influenced by the task itself, as observed in table 4: the use of “bus”, “hotel”, “international”, “conference”, “transport”, “accommodation”, etc. in the case of fourth year students and “lunch”, “dinner” and “hotel” in the first assignment written by first year students.

It is interesting to notice that the most frequent word in list three for both groups in the first assignment is “finally”, which seems to indicate a tendency of students to overuse some linking devices in detriment of other less known forms. There is also a coincidence of words in the second assignment. This time the most frequent adverbial form is “furthermore” (this shows some type of progression in their use of linking devices, but they still overuse some terms).

5.3. Textalyser

The results obtained for LD using two types of readability tests are illustrated in table 5:

Table 5. Results obtained using *Textalyser*

Groups	Lexical Density	Readability (Gunning-Fog Index): (6-easy 20-hard)	Readability (Alternative) beta: (100-easy 20-hard, optimal 60-70)
Corpus C	28.6%	11.7	35.8
Corpus A(C)	36.5%	6.2	67.3
Corpus A(D)	36.6%	6.4	68
Corpus B(C)	30.7%	10.6	40
Corpus B(D)	32.2%	10.4	41.9

As can be seen, the indexes of LD are lower when the first task in first year and the third are compared and they are even lower when compared with the results obtained for fourth year students. This, however, is understandable since LD has been reported by several authors to be affected by length, thus the LD index is lower in the case of fourth-year students.

In the results obtained for first-year students, there is a difference in the LD indexes of 2%–6% between the first and the second assignment. This may also be influenced by the nature of the task, and it indicates that the students may have used more lexical words in the first assignment and more functional words in the second one. In the case of readability, both tests demonstrate that the higher the level of the students is, the more complex their compositions are.

Textalyser also provides a list of the most frequent words in terms of occurrence, frequency and rank (table 6). The type of vocabulary employed by students in the different corpora indicates that some of the most frequent words in English are also found to be the most frequent in the three corpora. C2 students' most frequent words are "city" and "Valencia" and B1 students' in writing task 1 are "friends", "weekends", "see" and "tell". As for writing task 3, "English", "university" and "like" are the most frequent ones. Once more, we can notice the use of the intensifier "really", which appears as the third most frequent word in corpus A, showing, perhaps that our students overuse certain adverbs used for emphasizing while failing to use a wider variety of this type of intensifiers.

Table 6. Results on most frequent words drawn from *Textalyser*

Groups	Word	Occurrences	Frequency	Rank
Corpus C	city	207	3.8%	1
	you	127	2.3%	2
	valencia	82	1.5%	3
	our	77	1.4%	4
	your	57	1.1%	5
Corpus A(C)	you	98	3.3%	1
	went	85	2.9%	2
	really	50	1.7%	3
	friends	39	1.3%	4
	weekend	36	1.2%	5
Corpus A(D)	you	113	3.6%	1
	went	80	2.5%	2
	really	41	1.3%	3
	see	35	1.1%	4
	tell	35	1.1%	4
Corpus B(C)	english	92	3%	1
	university	72	2.4%	2
	course	54	1.8%	3
	like	51	1.7%	4
	you	41	1.4%	5
Corpus B(D)	english	98	3.2%	1
	university	84	2.8%	2
	like	47	1.5%	3
	your	42	1.4%	4
	you	34	1.1%	5

6. Conclusion

The study shows that it is possible to obtain a reliable measure of lexical richness which is stable across two pieces of writing by the same learners. It also demonstrates that we can discriminate between learners of different proficiency levels. For learners of English as a second language, the LFP is seen as a measure of how vocabulary size is reflected in use. In this study, LFP seems to correlate well with an independent measure of vocabulary size. This reliable and valid measure of lexical richness in writing will be useful for determining the factors that affect judgments of quality in writing and will be useful for examining how vocabulary growth is related to vocabulary use.

Students' written production has always been a central part in the assessment of student's linguistic competence. We believe that knowledge of lexical richness obtained through reliable quantitative and qualitative measures, like the ones we have looked at, may provide teachers with a more accurate picture of lexical progress, at least in the case of whole groups. Greater awareness of lexis on the part of the teachers is essential since it may help them to reflect on their teaching practice and on the suitability of their teaching materials. In spite of the difficulties of implementation, lexical richness should be a relevant factor to be taken into consideration in material design and learner assessment. The LFP measure can also be used as a diagnostic tool if teachers implement it to identify those students who rely excessively on high frequency words. Thus, this measure can become a pedagogic instrument to detect students who have problems due to poor productive vocabulary. Once those problems have been detected, it will be easier for teachers to help their students in the mastery of academic and low-frequency vocabulary.

Although this is a preliminary study, it seems that the LFP can provide interesting results but needs to be tested with a larger corpus and compared with other means that allow detection of progress in students such as vocabulary tests and teacher's assessment (Goodfellow et al 2002). We believe that studies like ours can shed light on the extent to which lexical richness can be a predictor of general essay quality (Lemmouh, 2008: 163)

Acknowledgements

The authors would like to thank Dr. Fuster-Márquez and Dr. Pennock-Speck for their valuable comments on all earlier versions of this study.

References

- Coxhead, A. (1998). An Academic Word List. *ELI Occasional Publications*, 18, School of Linguistics and Applied Language Studies, Victoria University of Wellington: Wellington.
- Coxhead, A. (2011). The Academic Word List ten years on: Research and teaching implications. *TESOL Quarterly*, 45/2, 355 - 362.
- Daller, H., van Hout, R., & Treffers-Daller, J. (2003). Lexical richness in the spontaneous speech of bilinguals. *Applied Linguistics*, 24/2, 197-222.
- Díez-Bedmar, M.B. (2011). Spanish pre-university students' use of English: CEA results from the University Entrance Exam. *International Journal of English Studies* 11/2, 141-158.
- Granger, S. (2002). A Bird's eye of corpus learner research. In Granger, S., Hung, J. & Petch-Tyson, S. (Eds.) *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching* (3-36), Amsterdam: John Benjamins.
- Granger, S. & Wynne, M. (1999). Optimising measures of lexical variation in EFL learner corpora. In J. Kirk (Ed.), *Corpora Galore* (249-257). Amsterdam and Atlanta: Rodopi.
- Goodfellow, R., Jones, G., & Lamy, M-N (2002). Assessing learners' writing using lexical frequency. *ReCALL vol 14 Part 1*. (133 – 145). Cambridge: Cambridge University Press.
- Hawkins, J.A. & Filipovic, L. (2012). *English Profile Studies 1. Criterial Features in L2 English. Specifying the Reference Levels of the Common European Framework*. Cambridge: Cambridge University Press.
- Halliday, M. A. K. (1985). *Spoken and written language*. Geelong Vict.: Deakin University.
- Heatley, A., Nation, I. S. P., & Coxhead, A. (2002). Range [Computer software]. Retrieved from <http://www.victoria.ac.nz/lals/staff/paul-nation/nation.aspx>
- Martin, S. & Hockey, H. (1988) Oxford Concordance Program. Oxford: Oxford University Computing Service.
- Nation, I. S. P., & Heatley, A. (1994). Range: A program for the analysis of vocabulary in texts [software]. Retrieved from <http://www.victoria.ac.nz/lals/staff/paul-nation/nation.aspx>

- Johansson, V. (2008). Lexical diversity and lexical density in speech and writing: a developmental perspective. Lund University, Dept. of *Linguistics and Phonetics Working Papers*, 53, 61-79.
- Laufer, B. and Nation P. (1995). Vocabulary Size and Use: Lexical Richness in L2 Written Production. *Applied Linguistics* 16/3, 307-322.
- Lemmouh, Z. (2008). The Relationship Between Grades and the Lexical Richness of Student Essays. *Nordic Journal of English Studies* 7/3, 163-180.
- MacWhinney, B. (1995). *The CHILDES Project: Tools for Analyzing Talk* (2nd ed). New Jersey: Lawrence Erlbaum.
- MacWhinney, B. & Snow, C. E. (1990). The Child Language Data Exchange System: an Update. *Journal of Child Language*, 17, 457-72.
- McCarthy, M. (1999) What constitutes a basic vocabulary for spoken communication? *Studies in English Language and Literature*, 1, 233-249.
- McCarthy, P., & Jarvis, S. (2007). Vocd: a theoretical and empirical evaluation. *Language Testing* 24, 459-488.
- McCarthy, P., & Jarvis, S. (2010). MTL-D, vocd-D, and HD-D: a validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods* 42, 381-392.
- McKee, G., Malvern, D., & Richards, B. (2000). Measuring vocabulary diversity using dedicated software. *Literary and Linguistic Computing*, 15, 323-337.
- Miller, J. F., & Iglesias, A. (2008). Systematic Analysis of Language Transcripts (SALT), English & Spanish (Version 9) [Computer software]. Madison: University of Wisconsin—Madison, Waisman Center, Language Analysis Laboratory.
- Textalyser.net TOOLS FOR ANALYSIS <http://textalyser.net/>
- Ure, J. (1971). Lexical density and register differentiation. In J.E. Perren & J.L.M. Trim (Eds.), *Applications of linguistics* (443–452). Cambridge: Cambridge University Press.
- Vermeer, A. (2000). Coming to grips with lexical richness in spontaneous speech data. *Language Testing*, 17/1, 65-83.
- Webb, S., & Paribakht, T.S. (2015). What is the relationship between the lexical profile of test items and performance on a standardized English proficiency test? *English for Specific Purposes*, 38, 34-43.