The Relationship Between Grades and the Lexical Richness of Student Essays

Zakaria Lemmouh, University of Stockholm

Abstract

The purpose of this study is to examine the relationship between lexical richness and the grades on essays produced by Swedish university students of English in order to shed light on the extent to which lexical richness is a predictor of overall essay quality. To this end, essays produced by 37 advanced learners of English were analyzed using a lexical richness measure that calculates the proportion of advanced vocabulary. The lexical richness score of the student essays were related to the following three variables: essay grade, course grade and vocabulary knowledge as measured by three discrete-item tests.

In addition, a 14-item questionnaire administered to the teachers at the English department eliciting information about their essay assessment procedures was analyzed, in order to shed light on the relationship between the weight teachers put on lexical richness and the grade they award essays with different lexical richness profiles. The results show that there is a relationship between use of advanced vocabulary in student essays and the overall course grade. However, no relationship was found between lexical richness and overall essay quality as reflected by faculty teachers' ratings. A possible explanation is that a majority of the surveyed faculty teachers state that their assessment of essay quality is primarily based on content and grammar features rather than lexical features.

1. Introduction

It is quite obvious that vocabulary knowledge and skills are important for successful communication in a second or foreign language. Words are the units of meaning which in turn make up sentences, paragraphs and entire texts. Corson (1997) underlines the necessity for learners in an academic context to gain productive written control of the Graeco-Latin vocabulary which dominates the vocabulary of the English academic language in order to be recognized as members of the academic writing community. A number of studies have demonstrated that a lack of vocabulary is what makes writing in a foreign language most difficult

(Uzawa and Cummings 1989, Raimes 1985, Leki and Carson 1994), and that vocabulary proficiency is perhaps the best indicator of overall text quality (e.g. Santos 1988; Astika 1993). In a number of studies different measures of lexical richness[1] in learners' written texts have been shown to relate well to the overall quality of the text (e.g. Jarvis 2002). The lexical richness of learners' written texts has been examined mainly from the point of view of the following five research questions:

1. Do measures of lexical richness provide consistent results when they are applied to two compositions written by the same learners, with only a short time interval in between?
2. How do the compositions of second language learners compare to those of native speakers of a similar age/educational level in terms lexical richness?
3. What is the relationship between lexical statistics and holistic ratings of learners' compositions?
4. What is the relationship between the lexical richness scores in learners' writing and their vocabulary knowledge as measured by a discrete-point vocabulary test?
5. Does the lexical richness of advanced learners' writing increase after one or two terms of English study?
(Read 2000:197-198)

The common denominator in these studies is that only one type of essay, viz. timed compositions, has been examined. When writing any particular essay type, the characteristics of that essay type result in statistical features specific to that text type, and different from other text types (McNeill 2007).

    Li (1997) investigated the extent to which lexical richness in EFL learners' timed compositions as measured by the *Lexical Frequency*

---

[1] Throughout this article the term 'lexical richness' will be used as a cover term for lexical statistics analyses of written texts, such as Lexical Originality, Lexical Density, Lexical Frequency Profile, etc.

Profile (LFP)[2] is related to teacher-raters' judgment of text quality. The findings show that teacher ratings and the LFP analysis seem to be able to discriminate between the best and weakest texts quite adequately, though less able to discriminate between average texts that are quite similar in nature.

Two other authors who use a different set of lexical richness analyses and who have also found a relationship between lexical richness and overall text quality are Linnarud (1986) and Engber (1995). The material analyzed in Linnarud (1986) consisted of 54 compositions written by 42 Swedish learners of English at the upper secondary level and 12 native speakers of English at the same age. The lexical richness measure for which she found a significant moderate correlation (0.47) with composition grades was lexical individuality, which measures the percentage of words in a composition that are unique to that specific composition in the entire text sample. Engber (1995) found a significant moderate correlation between lexical variation and scores of overall text quality of 0.57. She also carried out a count of lexical variation that included lexical errors but obtained a lower correlation score than the one without lexical errors. She employed an objective measure of overall text quality in the form of a 6-point scoring scale.

These findings suggest that lexical richness in learners' writing seems to be a moderately good predictor of overall text quality. The problem is that different studies obtain different correlations for different lexical measures. From the point of view of comparing results from different studies a universal measure of lexical quality would be ideal. The LFP has gained currency in later studies and is a measure that today comes closest to a standard analysis of lexical richness.

---

[2] The LFP shows the proportion of word families in an English target text coming from the four following word frequency levels: first 1000 most-frequent, second 1000 most-frequent, 570 most-frequent academic words and a low-frequency word level made up of words not contained in the previous three levels. The actual computation is carried out by a computer program termed *RANGE*, which can be downloaded from Paul Nation's web site: http://www.vuw.ac.nz/lals/staff/paul_nation/index.html.

In a more recent study by Morris and Cobb (2003), examining LFP as a predictor of academic performance among 151 TESL trainees with different language backgrounds, it was found that a significant low correlation holds between the proportion of academic words in the informants' texts and course grades. This correlation, is according to Morris and Cobb (2003), too low to warrant the use of LFP as the only assessment instrument of potential TESL candidates. However, the LFP was found to be able to discriminate between different proficiency levels. Since only texts of those accepted to the TESL programme were analyzed, what was examined was the degree of academic success of already successful applicants. If one extends the analysis to those who were refused entrance and to students on a similar programme with lower linguistic and academic standards, there is a statistically significant difference between these students and those accepted in the programme in terms of the proportion of advanced vocabulary in their timed composition.

All the above-mentioned studies have examined timed compositions from the point of view of the relationship between lexical richness and overall text quality or academic success. The relationship that has been found between lexical richness in timed compositions and different measures of language proficiency may differ according to writing task and the extent to which lexical richness is salient in the particular assessment adopted by the rater. As suggested by Engber (1995), the skills that are called upon in a timed writing task are different from those used in a process writing assignment, such as at-home essays. Kenworthy (2006) and Muncie (2002) are two studies which compare the lexical richness of timed compositions and at-home essays.

In Kenworthy, (2006) 16 university applicants whose L1 was Chinese wrote one timed composition and one at-home essay as part of the admittance requirements to an American university. A range of lexical features and grammatical errors were examined. The lexical features included the number of *cohesive devices, articles, pronouns, result clauses, adjective clauses, adverb phrases, prepositional phrases, synonyms, antonyms* and *demonstratives*. He found insignificant differences between frequency counts of lexical features in the timed compositions and in the at-home essays with the exception of synonyms. As regards grammatical errors, there were significant differences in the number of errors between the timed composition and the at-home essay,

leading him to conclude that the at-home essay writing task with its benefits of additional time and access to aids, positively affects overall textual quality. This only seems to apply to grammatical errors, but one can argue that the increase of the number of synonyms in the at-home essay indicates that at least this aspect of vocabulary use, which can be argued to be linked to lexical richness, also benefits from this writing format.

Muncie (2002), on the other hand, looked specifically at lexical richness as measured by the LFP when comparing timed compositions with a first and final draft of an at-home essay on the same subject. These were written by 30 Japanese EFL learners enrolled on an English composition course. She found that LFP scores did not improve significantly from the first to the last draft and that the last draft did not contain more Beyond 2000 vocabulary (for further details on Beyond 2000, see the method section) than the timed composition. However, when excluding final draft essays that did not bear any resemblance to the first draft, final drafts have an average Beyond 2000 score of 11.74% as opposed to 8.02% in the timed compositions. She states that

> if the timed composition is a measurement of their normal vocabulary range, the composition using the process approach shows not just their everyday range of vocabulary, but also the extra work and extra resources that the students have been able to employ during its production. (Muncie 2002: 232)

As is evident from Kenworthy (2006) and Muncie (2002), at-home essays differ from timed compositions in terms of grammatical and lexical richness, due to being written under different circumstances. Moreover, to my knowledge, all studies that have examined the relationship between lexical richness and some other measure of overall text quality or language proficiency have only examined one type of writing, namely timed compositions. Owing to the difference between timed compositions and at-home essays in terms of aspects such as different time constraints and language and writing skills involved, it is necessary to examine the relationship between lexical richness and measures of overall quality in at-home essays as well. Moreover, from a Swedish academic context it might be more relevant to examine at-home essays in this regard, since this type of writing assignment is a fundamental part of the course requirements in many Swedish universities. The first aim of the present study is to examine the extent to

which lexical richness in the particular writing task of at-home essays is related to teachers' holistic assessment as reflected by grades. A second aim is to investigate whether other variables such as course grade, which can be said to mirror overall language proficiency, are related to the amount of advanced vocabulary in student writing. Lastly, a third aim is to explore the degree to which faculty teachers lay emphasis on lexical richness in their assessment of student essays.

2. Method

The study reported in this article is part of my doctoral dissertation conducted at the English department at Stockholm University.

The informants are 37 first-term students at the English language and literature course at Stockholm University. Three vocabulary tests measuring different aspects of vocabulary knowledge were administered at the beginning and end of the term. These tests consist of two tests measuring receptive and productive size of vocabulary knowledge and a third test, developed specifically for the present study which measures productive depth of vocabulary knowledge. Whereas breadth of vocabulary knowledge is defined as the size of a learner's vocabulary, i.e. the number of words for which a learner can demonstrate at least a minimum of knowledge of meaning, depth is supposed to reflect how well various aspects of a word are known (Qian 2002).

The receptive vocabulary knowledge test was developed by Schmitt, Schmitt and Clapham (2001). The test involves word-definition matching. Each item consists of three definitions and six words. The productive size of vocabulary knowledge test was designed by Laufer and Nation (1999) and consists of 18 sentences per frequency level with a gap for which the test taker is prompted to supply the correct word.

It should be mentioned that these two standard tests as well as the LFP are premised on the idea that learners' vocabulary acquisition occurs in relation to the frequency of occurrence of words. In other words, high-frequency words tend to be acquired before low-frequency words. This idea is reflected in the measures by the incorporation of word frequency levels. The two discrete-items tests are made up of four levels of word frequency in English and an academic word level, viz. the first 2000 words, 3000 words, 5000 words, the Academic Word Level and 10,000

words. The Academic word level consists of words sampled from the Academic Wordlist (AWL; Coxhead 2000). The list contains 570 academic word families which occur frequently in a wide range of academic texts. The list does not include words that belong to the 2000 most frequent words of English. The depth of vocabulary knowledge test consists of 20 academic words sampled from the AWL. For each test item, the test-takers are prompted to supply a correct collocation, all the possible word class derivations and two synonyms.

In addition, all the essays that were part of the examination were collected and subjected to an LFP analysis which was designed by Laufer and Nation (1995). According to Laufer and Nation (1995), this measure overcomes various shortcomings of conventional lexical statistics. The LFP shows the relative proportion of words from different frequency levels in a written text. The LFP calculates the proportion of words that belong to the following four levels or lists: the first 1000 most frequent words, the second 1000 most frequent words, the AWL level and a fourth level called the 'not-in-the-lists' word list consisting of words not contained in any of the other levels. In the present study a condensed profile called the Beyond 2000 (B2000) measure will be used. It calculates the proportion of words not contained in the first and second 1000 most frequent word levels. In other words, a B2000 profile is simply attained by adding the last two levels, namely the AWL and the not-in-the-lists word list. The underlying idea behind this measure is that the higher proportion of B2000 words, the higher the lexical richness of the text (c.f. Laufer 1995).

All subjects agreed to hand in all essays they completed as part of the course requirements. I also had informants' permission to inquire about their grades. The informants are in two groups: one group consisting of 17 students who started their studies in the fall of 2006 and a second group of 20 students who started their first term in the spring of 2007. Both groups were graded according to the old grading system in which the grades are as follows: Fail, Pass (P) and Pass with distinction (PwD). All students were required to write at least three literature essays in the form of a close reading within the three genres of *poetry*, *fiction* and *drama*, and a linguistics essay on a topic of their choosing.

As regards the first group of students, with the exception of three informants, only a fourth final close-reading was graded. In the second group of students, most of them only wrote three close readings that all

were graded; four informants wrote four close-readings, of which only the fourth final close-reading was graded. The number of essays the students had graded varied according to who their seminar teacher was. In other words, the teachers adopted different grading routines. The linguistic essays for both groups were graded.

Before entering the essays into the computer program *Range*, in which the LFP statistics are calculated, all words that were clearly used incorrectly were omitted, as they could not be considered as part of the learner's productive vocabulary. This did not occur often. If, on the other hand, a word was used correctly but misspelled, it was corrected and retained. A wrong derivative of a word was not considered wrong since all derivatives that make up one word family have the same frequency. Proper names were omitted from the sample since they are not covered by the frequency levels (c.f. Laufer and Nation 1995). Moreover, nouns and adjectives denoting nationality were omitted since frequent use of less frequent nationality words might skew the profile.

In order to elicit information on the degree to which teachers focus on lexical richness in their assessment of student essays, a 14-item questionnaire was administered to 10 linguistics and literature teachers at the department.

There are about 20 teachers currently working at the department. The number varies slightly from term to term. The criterion for being requested to participate in the survey was that the teachers within the last two years had graded essays at the first level, which corresponds to the first term of study in the English language and linguistics and literature programme. Out of the 15 teachers that fulfilled this criterion ten completed the questionnaire. The sample consists of five linguistics and five literature teachers.

3. Results

In this section the results of the present study will be outlined. The results pertaining to the student essays will be presented first, followed by a presentation of the teacher questionnaire data.

3.1 The relationship between lexical richness and essay and course grade

Table 1 below compares the average proportion of B2000 words in all the essays of students awarded the course grade Pass (P) and students awarded the grade Pass with Distinction (PwD).

Table 1. Comparison of the average proportion of B2000 words in all essays produced by students awarded the course grade P and PwD

Course grade	B 2000	Mean difference	St.d.
P (n=15)	7.82%	1.55%*	1.47
PwD (n=12)	9.37%		1.93

*Correlation is significant at the 0.05 level (2-tailed).

In order to determine whether there is a statistically significant difference between the groups, an independent-samples T-test was carried out. It can be seen in Table 1 that there is a statistically significant difference of 1.47% between P students and PwD students.

The difference between P and PwD students seems to suggest that there is a link between using relatively more advanced words and achieving a relatively higher academic success in English studies. Moreover, the LFP measure seems to be able to discriminate between different proficiency levels as reflected in the awarded course grade.

Table 2 below displays the results from an independent-samples T-test comparing the proportion of advanced vocabulary in literary essays awarded the grade P and PwD.

Table 2. Comparison of the proportion of B2000 words in literary essays awarded the grade P and PwD

Average essay grade on literary essays	B 2000	Mean difference	St.d.
P (n=20)	8.07	0.49	1.58
PwD (n=13)	8.56		2.21

There is a small, non-significant mean difference in the average proportion of B2000 words between literary essays awarded the grade P and essays awarded the grade PwD. Accordingly, these results seem to suggest that essay quality in terms of the awarded essay grade is not related to the proportion of advanced vocabulary in student writing. In order to investigate whether this might be due to a genre effect, in that

the degree of advanced vocabulary is not a crucial part of the criteria for a literary essay to be awarded the grade PwD, an analysis of linguistics essays along these lines was carried out.

Table 3 below shows the mean difference in the proportion of B2000 in the linguistics essay between essays awarded the grade P and essays awarded the grade PwD.

Table 3. Comparison of the proportion of B2000 words in the linguistics essays awarded the grade P and PwD

Average essay grade on linguistics essays	B 2000	Mean difference	St.d.
P (n=17)	10.1935	-0.17	2.56
PwD (n=13)	10.0188		2.64

As can be seen in Table 3 there is no statistically significant difference in the proportion of B2000 words in the linguistics essays awarded the grade P and those awarded the grade PwD. The LFP measure does not seem to discriminate between essays in terms of the awarded grade. Moreover, there does not seem to be a difference between literature and linguistics teachers as regards the weight they put on the proportion of advanced vocabulary in student writing.

In order to investigate to what extent vocabulary knowledge is associated with lexical richness in student essays, a correlation analysis was carried out. Table 4 below presents the Pearson product-moment correlations between, on the one hand, the average proportion of advanced vocabulary in all student essays and, on the other hand, the scores of the receptive vocabulary levels test (RVLT), the productive levels test (PVLT) and the depth of vocabulary knowledge test (Depth). The rationale for this analysis is that it might shed light on the extent to which learners' vocabulary knowledge accounts for the degree of lexical richness in their essays. In contrast to timed composition tasks in which the writer can only rely on his actual language and writing ability, the quality of the at-home essay might to a higher degree reflect a writer's ability to use external aids rather than the writer's language and writing proficiency.

Table 4. Pearson correlations between B 2000 and Rec. VLT, Prod. VLT and Depth

	RVLT	PVLT	Depth
B2000 all essays (n=37)	.231	.408*	.420**

*Correlation is significant at the 0.05 level (2-tailed).
** Correlation is significant at the 0.01 level (2-tailed).

As shown in Table 4 above, the three tests form an ascending order in terms of the degree to which they correlate with the B2000 score. As one might expect, the RVLT scores do not correlate significantly with the B2000 scores. This is probably due to the fact that the RVLT measures a receptive ability whereas the B2000 score reflects a productive ability. Both the productive tests (PVLT and Depth) show a statistically significant correlation with the B2000 score.

These results seem to suggest that those students who have a large productive vocabulary and extensive knowledge of academic words tend to use more advanced vocabulary in their writing.

Table 5 below shows the mean difference in the score obtained on the three vocabulary tests between students awarded the course grade P and students awarded the course grade PwD.

Table 5. Comparison of vocabulary test scores of students awarded the course grade P and students awarded the course grade PwD

		N	Mean	Mean difference	St.d.
RVLT	P	15	264.13	10.20*	11.74
	PwD.	12	274.33		10.33
PVLT	P	15	109.33	14.42*	15.21
	PwD.	12	123.75		11.51
Depth	P	15	49.07	14.43*	15.86
	PwD.	12	63.50		11.82

*Correlation is significant at the 0.05 level (2-tailed).

As can be seen in Table 5, there is a statistically significant difference in the vocabulary test scores across the board between P students and PwD students. On average the PwD students obtained a 4% higher score on the RVLT than the P students. Although statistically significant, it is relatively low. However, as regards the two productive vocabulary tests, the PwD students obtained on average a 12% higher score on the PVLT and about a 23% higher score on the Depth test. The relatively low

difference between P students and PwD students in their RVLT scores might be due to a ceiling effect.

From the data in Tables 4 and 5 we can see that, although the students might use different aids such as dictionaries when writing at-home essays to enhance lexical richness, actual vocabulary knowledge seems to account for a large portion of the lexical richness in their writing, thus lending concurrent validity to the LFP measure.

3.2 Teacher responses to a questionnaire on student essay assessment

The degree to which any language feature such as vocabulary or grammar is related to essay quality in terms of teacher ratings, depends mainly on two aspects: which features teachers focus on when assessing student writing and the extent to which the raters have adopted an agreed upon objective standard for rating student texts. The first aspect has to do with the degree to which any one feature is focused on by the teachers in a consistent manner. In regard to the second aspect, a high degree of inconsistency in which language or content features are emphasized by the teachers will make it difficult to find a pattern in terms of the relationship between, in this case, the proportion of advanced vocabulary and grades.

The following question was intended to elicit information on whether the teachers employ any written criteria when rating essays: *Do you follow any set of written criteria when rating essays?* The responses showed that only three out of the ten respondents use any written criteria when rating essays. Although use of objective standards when rating essays does not guarantee reliability in terms of intra- or inter-rater reliability, it does at least increase the degree of consistency in both these regards. Leki (1995: 24) states that:

> That we share standards and expectations of 'good writing' is implicit in our teaching and assessment of writing. But the problem with these standards and expectations is that we cannot be certain if, or to what degree, our assumptions are shared by other constituents of our community.

Accordingly, one reason for there not being a significant difference between P and PwD essays in terms of the proportion of advanced vocabulary might be due to a high degree of inconsistency in the way the teachers approach different language features in the essays.

As for how much weight the respondents put on content features in comparison to language features in their assessment of student essays, four of the respondents state that they put equal weight on language and content features. Four of the respondents state that they put more weight on language features than content. Two of the respondents state that they focus more on content than language features when rating essays. Accordingly, a majority of the teachers put equal or more weight on language features than content. This seems to suggest that language proficiency plays an important role in how essays are graded. However, a majority of the respondents (n=7) stated that within language they put more weight on grammar than lexical features. In order to shed light on the degree to which the respondents view use of advanced vocabulary as a major vocabulary feature of a good essay, the respondents were asked to list the vocabulary features they considered as indicative of a good essay.

Figure 1 illustrates what lexical features the respondents regard as important vocabulary features of a good essay. [the information in this sentence is repeated in the next three/four sentences; perhaps you could reformulate or shorten]
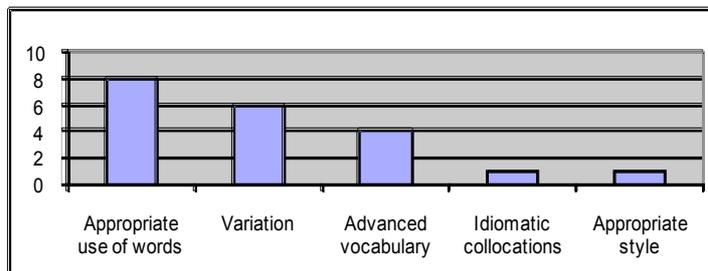

Figure 1. The important vocabulary features of a good essay

The vocabulary features that the respondents listed as characteristic of a good essay fall into five separate categories of which *appropriate use of words* is the feature that most frequently was reported as an important vocabulary feature of a good essay. The second most frequently listed vocabulary feature was *variation* which was reported by six of the respondents as characterizing a good essay. Use of *advanced vocabulary* was reported by four of the respondents as a important vocabulary

feature of a good essay by four of the respondents. The last two vocabulary features, use of *idiomatic collocations* and *appropriate style* were listed as important vocabulary features of a good essay by one respondent each.

Appropriate use of words is mainly concerned with depth of vocabulary knowledge in that it is not sufficient for learners to have superficial knowledge of a word; they must also have knowledge of a word's range of meanings and register constraints in order for the use of a specific word to be considered appropriate by the rater. This specific aspect of vocabulary use is not reflected in the LFP analysis to a very great degree since only words that are clearly used wrongly are discarded. Thus, words are retained that, although not clearly used incorrectly, might be assessed by the teachers as inappropriately used, and these words might then contribute to a negative assessment of the quality of the essay at hand.

The second most frequently reported vocabulary feature, *variation*, is not reflected in the LFP analysis, since the LFP does not calculate lexical variation. The third feature, *advanced vocabulary*, on the other hand, is measured by the LFP. Based on the responses to this question one can draw the conclusion that the two most frequently reported vocabulary features regarded as characterizing a good essay are not captured by the LFP analysis. This might be a factor in the degree to which the proportion of B2000 vocabulary is related to the teachers' assessment and, by extension, to the grade given to a specific essay.

Let us now turn to three hypothetical questions aimed at reflecting the degree to which the respondents relate grammar and vocabulary to text quality.

Figure 2 below shows the number of respondents who answered yes or no to the following question: *Can you have a good essay with poor grammar*?
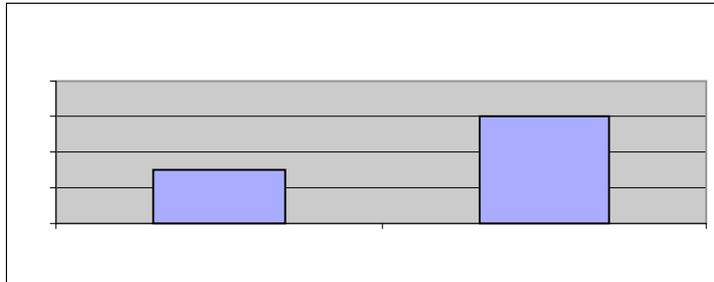
Figure 2. Can you have a good essay with poor grammar?

Nine of the ten respondents answered this question, and the majority (n = 6) answered <u>no</u> to it.

In response to the following two questions :

- *Can you have a good essay with poor vocabulary, e.g low degree of lexical variation and high dependence on high frequency words?*

- *Can you have good vocabulary but a weak essay, e.g. a high degree of lexical variation and a low dependence on high frequency words?*

six of the teachers answered <u>yes</u> to the first question and eight of the respondents answered <u>yes</u> to the second question. Accordingly, extrapolating from this, a majority of the respondents do not regard good vocabulary as a decisive factor in the overall quality of an essay.

Relating these results to the data displayed in Figure 2, it can be surmised that in terms of overall quality correct grammar seems to be a more crucial factor than good vocabulary. How this is actually manifested in practice in the teachers' ratings of essays is beyond the scope of this study. Suffice it to say that this general point of view among the teachers surveyed in the present study will affect the degree to which a high proportion of advanced vocabulary in student essays is predicative of a higher grade.

4. Conclusions

There seem to be contradictory results regarding the relationship between lexical richness and grades. On the one hand there is a relationship between course grade and the average proportion of advanced vocabulary in student essays; on the other hand, there is no significant difference in terms of the proportion of B2000 words between essays awarded the grade P and PwD.

However, when one takes into account the teachers' rating practices it becomes evident that lexical richness is only one of many features the teachers include in their assessment. The LFP measure does not seem to be a good measure of overall text quality when it comes to at-home essays, regardless of genre, simply because the assessment of at-home essays seems to be more content and grammar oriented.

Although the LFP does not discriminate between at-home essays in terms of their overall quality as reflected by the awarded grade, it seems to be a valid predictor of academic success. The LFP has been shown to be related to both productive size of vocabulary knowledge and productive depth of vocabulary knowledge. Moreover, it has been shown to discriminate between different proficiency? levels as reflected by course grade.

As regards earlier studies that have shown that lexical richness is related to the overall quality of compositions, one plausible reason for earlier research having found such a relationship might be that the holistic ratings to which the lexical richness scores were correlated to, emphasise? lexical features, such as lexical richness and quality. [the following sentence is incomplete]: Hence, the relatively strong correlations found between lexical richness and the adopted holistic ratings. In light of this, when examining the relationship between linguistic features and the overall quality of a written text, it is important to delineate the criteria in the holistic rating, to which the examined linguistic features are related.

From a pedagogical point of view two main conclusions can be drawn from the findings. Firstly, the LFP measure can be used as a diagnostic tool to identify students who in their writing predominantly rely on high-frequency vocabulary. It can thus be used as pedagogical tool to spot learners early on who run the risk of failing the course due to a poor productive vocabulary. Secondly, if indeed the proportion of

advanced vocabulary in students' at-home essays is related to language proficiency, lexical richness should be emphasized in the evaluation of student texts to a greater extent. A number of studies have shown that the mastery of academic and low-frequency vocabulary is strongly related to academic success (e.g. Nation 2001, Laufer 1998, Jarvis 2002). In the present study it has been shown that students awarded the course grade PwD produce more academic and low-frequency vocabulary in their writing. In the light of these findings one might argue that vocabulary features should receive more focus in the teacher assessments of essays produced by Swedish university students of English, since it might further encourage the students to improve their vocabulary knowledge.

References

Astika,Gusti Gede.1993. "Analytical assessments of foreign students' writing." *RELC Journal* 24**:** 61-71.
Corson, David. 1997. The learning and use of academic English words. *Language Learning* 47: 671-718.
Coxhead, Averil. 2000. "A new academic word list." *TESOL Quarterly* 34: 213-238.
Engber, Cheryl. 1995. "The relationship of lexical proficiency to the quality of ESL compositions." *Journal of Second Language Writing* 4: 139-155.
Jarvis, Scott. 2002. "Short texts, best-fitting curves and new measures of lexical diversity." *Language Testing* 19: 57-84
Kenworthy, Roger. 2006. "Timed versus at-home assessment tests: Does time affect the quality of second language learners' written compositions?" *TESL-EJ* 10: 10-22.
Laufer, Batia. 1995. "Beyond 2000: a measure of productive lexicon in a second language." *The Current State of Interlanguage*. Eds. Lynn Eubank, Larry Selinker and Michael Sharwood Smith. Amsterdam**:** John Benjamins Publishing Company. 265-272.
Laufer, Batia. 1998. "The development of passive and active vocabulary in a second language: Same or different?" *Applied Linguistics* 19: 255-271.
Laufer, Batia and Paul Nation. 1995. "Vocabulary size and use: lexical richness in L2 written production." *Applied Linguistics* 16: 307-322.

Laufer, Batia and Paul Nation. 1999. "A vocabulary-size test of controlled productive ability."*Language Testing* 16: 33-51.

Leki, Ilona and J. Carson. 1994. "Students' perception of EAP writing instruction and writing needs across the disciplines." *TESOL Quarterly* 28: 81-101.

Leki, Ilona. 1995. "Good writing: I know it when I see it." *Academic writing in a second language*. Eds. Diane Belcher and George Braine. New Jersey: Ablex Publishing Corporation. 23-46.

Li, Carrie. 1997. *Is lexical richness an essential criterion in judging a piece of writing?* Master's dissertation presented at the University of Hong Kong.

Linnarud, Moira. 1986. *Lexis in composition: A performance analysis of Swedish learner's written English*. Malmö: CWK Gleerup.

McNeill, Brian. 2007. "A summary of 'A comparative statistical assessment of different types of writing by Japanese EFL college students." *Journal of School of Foreign Languages* 33: 141-155.

Morris, Lori and Tom Cobb. 2003. "Vocabulary profiles as predictors of the academic performance of teaching English as a second language trainees." *System* 32: 75-87.

Muncie, James. 2002.Process writing and vocabulary development: comparing Lexical Frequency Profiles across drafts." *System* 30: 225-235.

Nation, Paul. 2001. *Learning vocabulary in another language*. Wellington: Cambridge University Press.

Qian, David. 2002. "Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective." *Language Learning* 52: 513-536.

Raimes, Ann. 1985. "What unskilled ESL students do as they write: a classroom study of composing." *TESOL Quarterly* 19: 229-258.

Read, John. 2000. *Assessing vocabulary*. Cambridge: CUP.

Santos, Terry. 1988. "Professors' reactions to the academic writing of nonnative-speaking students." *TESOL Quarterly* 22: 69-88.

Schmitt, Norbert, Diane Schmitt and Caroline Clapham. 2001. "Developing and exploring the behaviour of two new versions of the vocabulary levels test." *Language Testing* 18: 55-88.

Uzawa, Kozue and Alistair Cumming. 1989. "Writing strategies in Japanese as a foreign language: lowering or keeping up the standards." *The Canadian Modern Language Review* 46: 179-191.