

The relative significance of lexical richness and syntactic complexity as predictors of academic reading performance

Karami, Mehdi ✉

Azerbaijan Shahid Madani University, Iran (karami_m58@yahoo.com)

Salahshoor, Farzad

Azerbaijan Shahid Madani University, Iran (farzad.salahshoor@gmail.com)



ISSN: 2243-7754
Online ISSN: 2243-7762

Received: 2 June 2013

Available Online: 10 November 2013

Revised: 16 October 2013

DOI: 10.5861/ijrsl.2013.477

Accepted: 17 October 2013

OPEN ACCESS

Abstract

Reading academic texts that include mainly university textbooks has been a challenge for EAP learners. There are various reasons for text difficulty; however, linguistic elements were investigated in this study. The aim of this study was to determine whether lexical richness of the readers would be a more potent predictor of their academic reading performance or their ability for producing and processing complex syntactic structures. The study involved 50 ELT teacher trainees, 25 juniors and 25 seniors, at Shahid Madani University of Azerbaijan, Iran. In a standard multiple regression design, the participants were given an opinion essay-writing task and an IELTS academic reading test. Their scores on IELTS academic reading test were regressed against LFP (Lexical Frequency Profile) and MLTU (Mean Length of T-Unit) indexes of their essays. LFP index is a measure of lexical richness adapted to the web for free online access under the name Web-VocabProfile, and MLTU index is a measure of syntactic complexity. Results indicated that the ability in producing and processing complex syntactic structures rather than mere grammatical knowledge can be considered as effective a predictor of academic reading comprehension as lexical richness. Therefore, lexical richness may no longer be supposed as the single most important predictor of academic reading performance.

Keywords: academic reading performance; syntactic complexity; MLTU; lexical richness; Web-Vocab Profile

The relative significance of lexical richness and syntactic complexity as predictors of academic reading performance

1. Introduction

It is assumed that EAP learners who are to acquire knowledge through the medium of English in their universities should have a certain level of proficiency in the language. However, academic texts have different linguistic requirements from non-academic ones, since they belong to different genres, i.e., expository versus narrative, and each genre has its own features. They are different lexically as well as syntactically; therefore, EAP learners are supposed to learn those lexical items and grammatical patterns which are more frequent in academic texts. Scarcella (2003) argued that learners need to be familiar with lexical and syntactic features particular to a discipline. The lexical components of academic English include knowledge of sub-technical words in addition to general words (p. 12). Sub-technical words are common-core words frequent in academic texts. Nation and Waring (1997) emphasized, for academic purposes, a learner needs knowledge of AWL – Academic Word List – in addition to less-frequent words, but the need for GSL – West’s (1953) General Service List which involves the first 2000 most frequent words in English – is less. Scarcella (2003) also asserts that, in academic English, students need knowledge of particular syntactic features other than those necessary for ordinary English. They need knowledge of additional structures such as complex sentences, parallel clauses, conditionals, and passive sentences (p. 12). Complex structures have always had a higher profile in academic texts than in general texts like fiction and conversation, since, in academic register, ideas need to be linked together to create a cohesive discourse (Conrad, 1999; Dudley-Evans & St John, 2005).

1.1 The significance of the study

This study aimed at investigating the relative significance of lexical richness and syntactic complexity in academic reading comprehension. A substantial amount of empirical research verifies that lexical knowledge is the most important competence for academic needs (Qian, 2002; Cobb & Morris, 2004). Neither of these studies examined the impact of syntactic complexity or its interaction with lexical complexity in academic language. Moreover, many of these studies have the main drawback of measurement instruments. Academic texts have certain lexical items frequent in them (academic or sub-technical words) which knowledge of them seems to be a determining factor for better comprehension, while measures of lexical knowledge in the previous studies did not include these words in their vocabulary tests. This problem was also prevalent in measures of syntactic knowledge. Academic texts employ excessive use of complex structures like subordinations, the mere knowledge of which as grammatical rules alone is not sufficient for efficient comprehension of such texts. What is problematic with these structures is their processing which is a large burden on working memory. Therefore, there should be measures which can test participants’ ability in producing and processing them, while measures in the previous studies (Ransdell, 2003; van Gelderen et al., 2003; Shiotsu & Weir, 2007) tested only declarative knowledge of grammatical rules. Therefore, the researcher aimed at avoiding such failures by employing measures of lexical richness and syntactic complexity.

1.2 The purpose of the study

The purpose of the study was to find out whether EAP learners’ lexical richness – use of vocabulary in context – is a more potent predictor of their academic reading performance or their ability in producing complex syntactic structures. The results of the study may contribute to the syllabus designers, teachers, and learners in EAP fields who want to deal better with linguistic features of academic texts. They may suggest a broader hypothesis for further research into the significance of syntactic complexity in reading academic texts.

1.3 Academic reading

Uribe (2008) states academic English is needed by the educated for the university studies and not for everyday situations (p. 1). Therefore, academic reading is a skill necessary for EAP or EFL learners so that they are enabled to study semi-technical and technical essays, articles, and university textbooks. In testing academic reading, there are a few standardized tests like IELTS academic reading module, UETESOL (University Entrance Test in English for Speakers of Other Languages), and MELIAET (Michigan English Language Institute's Academic English Test) (Hamp-Lyons, 2002, p. 129; Dudley-Evans & St John, 2005, p. 214).

1.4 Lexical richness

Most of researchers have agreed upon vocabulary use as a component of lexical ability (Nation, 1990, p. 31; Paribakht & Wesche, 1993; Chapelle, 1994, as cited in Read, 2000, p. 28; Henriksen, 1999; Carter, 2000, p. 239). Recently, lexical richness, also known as 'vocabulary use in context' (Read, 2000) or 'vocabulary size reflected in use' (Laufer & Nation, 1995), has stimulated some interest as an aspect of vocabulary knowledge. Read and Chapelle (2001) noted that most existing vocabulary tests had a trait view of vocabulary knowledge, i.e., these measures implicitly advocated the notion that as a mental attribute of the learner, word knowledge could be described and measured without any reference to the contexts in which the words are used. In an effort to avoid this failure, Read and Chapelle (Read, 2000) called for the development of tests which would assess the learners' ability to deploy their vocabulary knowledge in particular contexts of use, since interpretation of a word and knowledge of that word is influenced by the context it is used in (p. 4).

In line with this development, Laufer and Nation (1995) devised a new measure which they claim tests the learners' lexical richness. Lexical Frequency Profile (LFP) is based on the relative frequency of words in the language and involves simply calculating the percentage of word families in the learners' composition that belong to each of three or four frequency bands. In their original study, Laufer and Nation used a profile that comprised four levels: first 1000 most frequent words (K1) and second 1000 most frequent words (K2) based on West's (1953) GSL, words in the Academic Word List (AWL) based on Coxhead's (2000) AWL, and other less-frequent words outside these two lists of words (Off-List or OL). Thus, if a learner wrote an essay containing 200 word families in total and these consisted of 150 from K1, 20 from K2, 20 from AWL, and 10 from OL, the LFP index for it would be 75%-10%-10%-5% respectively. Cobb and Morris (2004) in analyzing comparative results of native and non-native speakers' vocabulary profiles established a standard for LFP indexes by correlating the levels which resulted systematically in academic success. The standard levels arrived at were the following: a K1 score of less than 85%, an AWL score of over 5%. Cobb (2009a) adapted this LFP program to the web for free online access under the name Web-VocabProfile (Web-VP). Once a text has been typed into the program window, VP determines what percentage of the words in the text is covered by each of K1, K2, AWL, and OL.

Read and Chapelle (2001) have argued about approaches to construct definition in vocabulary assessment. There are three definitions, and LFP is placed in the most comprehensive one, the interactionist approach, which requires researchers to specify the relevant aspects of both trait and context because it refers to a context-specific underlying ability. LFP is such a measure which assesses learners' ability to use vocabulary correctly. Read's (2000) argument is that LFP provides a more objective and differentiated measure of the learners' vocabulary use and lexical richness than the other statistics because it draws on a range of established word lists (GSL and AWL) to classify the words into categories.

1.5 Syntactic complexity

All the present models of communicative language ability (Canale & Swain, 1980; Bachman, 1990; Celce-Murcia et al., 1995; Bachman & Palmer, 1996) have included syntactic knowledge as one sub-component necessary for successful communication. Besides accuracy and range, structural complexity of the written or

spoken discourse produced by the L2 learner is a measure of his / her syntactic knowledge. Ortega (2003) stated that “syntactic complexity refers to the range of forms that surface in language production and degree of sophistication of such forms” (p. 492). Carroll (2008) defines syntactic complexity as “the complexity of the expressions used to convey the idea” (p. 288). On the assumption that as language learners progress, their grammar becomes more complex, the degree of complexity would seem to be a necessary criterion in construct validation (Rimmer, 2006; Ortega, 2003). Mousavi (1999) has noted that syntactic complexity is used to examine the oral and written language of the individuals to determine their syntactic growth (p. 384). Therefore, this is not accurate knowledge, alone, of grammar that leads to success in communicative language ability, but complexity of that knowledge is also important. Length of production unit, amount of embedding, range of structural types, and sophistication of the particular structures deployed in production have all been the target of quantifications when characterizing syntactic complexity, resulting in a variety of global measures (Ortega, 2003, p. 492). The cumulative effect of these features is striking in complexity of structures.

Grammatically speaking, a complex sentence is a ‘sentence which contains one or more dependent clauses, in addition to its independent, or main clause (Richards & Schmidt, 2010). However, if we define a complex sentence psycholinguistically as one that expresses more than one proposition; passive sentences, complements, coordination, relative clauses, and subordinations are among complex sentences, since they convey a single idea in a linguistically complex form that is difficult to process (Carroll, 2008, p. 293). Therefore, sentences may be syntactically complex not simply for their grammatical definition, but because of complexity in processing them due to limitations in working memory capacity. Carpenter et al. (1994), Gass and Selinker (2008), and Carroll (2008) noted that the working memory load imposed by comprehension is directly related to sentence complexity. Long and embedded sentences as complex ones increase processing load and “this leads to deterioration of linguistic performance” (Smith & Tsimpli, 1995). Cook and Newson (1996) argued that memory capacity affects the length of a sentence that can be produced or understood but has nothing to do with knowledge of language itself (p. 26). Perfetti, Landi, and Oakhill (2005) also found that problems in parsing complex syntactic structures arise from processing and working memory limitations not simply from a lack of knowledge of grammatical rules. A sentence like “Canaries have wings.” demands few attentional resources whereas a sentence like “The point about photography – and I shouldn’t really be saying this because I write books about it – is that you have to keep at it.” demands more (Field, 2003, p. 113). Number of words per T-unit is a measure for syntactic complexity that has the obvious advantage that it is a straightforward and the most time-honored measure and not even necessarily involves manual coding (Szmrecsányi, 2004).

1.6 Mean length of T-unit

T-Unit or Minimal Terminable Unit is a measure of syntactic complexity of the sentences. Mean number of words per T-Unit or MLTU, accordingly, can be used as a measure of syntactic complexity of a text. It is a basic unit of expression which is defined as “a main clause plus all subordinate clauses and non-clausal structures attached to or embedded in it” (Mousavi, 1999, p. 430; Richards & Schmidt, 2010). A ‘simple’ or a ‘complex’ sentence contains just one T-Unit because they have only one main clause. For example, the sentence ‘The fire burned.’ is one T-Unit; so is ‘The fire burned until the fuel ran out.’ A ‘compound’ sentence is more than one T-Unit, since it contains more than one main clause. For example, ‘The fire burned, and the fuel supply dwindled.’ has two T-Units (Mousavi, 1999, p. 430).

Gass and Selinker (2008) have stated that with written data, error-free MLTU is a more precise measure than standardized tests and teacher evaluations (p. 74). MLTU as a measure of both syntactic and psychological complexity has psychological reality, since complex sentences require more processing abilities in both production and comprehension than compound sentences (Kess, 1992, pp. 321, 323). Ortega (2003), in a meta-analysis, reviewed 25 studies rating the syntactic complexity of the written work of EFL and ESL learners. The most frequent measure employed in these studies was MLTU. Homburg (1984) considered MLTU as a valid measure of syntactic complexity and writing quality and reported that MLTU along with other measures like ‘concept of error’ account for 84% of the variance in writing proficiency. Jiang (2013) also reported that

'percentage of error-free T-units is found to be the only measure that discriminates between all levels of a Chinese learner cohort.

2. Review of Literature

Cobb and Morris (2004) conducted a study in which they examined the potential offered by VocabProfile (LFP) indexes as predictors of academic performance in undergraduate TESL programs. Vocabulary profiles were established for 122 TESL students by means of an analysis of 300-word samples of their writing. The students' scores on each profile component (K1, K2, AWL, and OL) were then correlated with the grades they were awarded in two of the grammar courses (EAP courses titled 'How to teach grammar?') in their program of study. The findings of the study revealed that the students' vocabulary profile results, especially at AWL, correlated significantly, $r = .37$, $p < .01$, two-tailed, with grades in the more procedurally oriented of the two courses. Cobb and Morris's results are in concordance with Nuttal's (1996) idea that "[EAP] students often have difficulty not with the technical jargon but with the common-core of the semi-technical words [AWL] that occur in most disciplines" (p. 67). In verification of this idea, Lemmouh (2008) in a study has shown that students who were awarded the course grade "Pass with Distinction" produce more academic and low-frequency vocabulary in their writing (p. 179). Therefore, richness of the learners' free-active vocabularies, especially in AWL category, may be helpful in reading English for General Academic Purposes (EGAP) texts.

Alderson (2000) stated, "Understanding syntactic patterns is necessary for reaching 'effectiveness level' in academic reading (p. 135). Martohardjono et al. (2005) in their study attributed difficulties in the higher level text comprehension to 'Structural Deficit Hypothesis' (SDH) which is a deficiency in processing complex syntactic structures. They reported that it is a certain type of syntactic knowledge which is important for comprehension; specifically, it is the ability to process complex syntactic structures, like subordinate clauses. The scores of the subordination part of their grammar test correlated highly with reading comprehension ($r = .57$, $p = .038$). Cohen et al. (1979) conducted a study titled 'What is problematic for non-native readers when they are reading materials in a specialized field written in English?'. Their subjects were 12 university students majoring in biochemistry, biology, economics, and history who acted as informants. In all four studies, the students were asked to underline all the words and structures they found difficult to understand in an extract from their course readings; they were also asked to answer content questions. Across disciplines, the students were found to have similar problems. These included complex subordinated structures and sub-technical words. Geva (1992) reported that university students knowledge of logical connectors (subordinators) had a significant correlation with Academic Text Comprehension (ATC) scores, i.e., $r = .40$, $p < .001$ in SCT (Sentence Completion Task test) at sentential level, and $r = .49$, $p < .001$ in cloze measure at discourse level. The results indicated that knowledge of complex structure markers that form T-Units has a significant correlation with higher-level academic text comprehension.

2.1 Research question

The current study intends to address the following question: Which is the best predictor of academic reading performance: syntactic complexity or lexical richness?

3. Methodology

3.1 Participants

The study involved 50 majors of Teaching English as a Foreign Language, 25 juniors and 25 seniors, registered in 4-year BA programs at Iranian Azerbaijan University of Teacher Training (Shahid Madani University). They were all native speakers of Persian who were learning English as a foreign language. Thirty male and twenty female students at the age range of 18 to 24 of four intact classes were available at a certain time to the researcher, therefore, the researcher employed convenience or opportunity sampling technique.

3.2 Materials

IELTS academic reading tests were employed to test the participants' academic reading ability. VocabProfiler (VP) was used as a measure of lexical richness. Mean Length of T-unit (MLTU) was also employed as a measure of syntactic complexity.

To test the reliability of the participants' scores in academic reading tests of IELTS, Cronbach's alpha was calculated. IELTS academic reading test proved to be highly reliable in an internal reliability test of 'Cronbach Alpha', $\alpha = .81$. IELTS academic reading test could also differentiate between levels of proficiency. The mean difference was largely significant, positive for seniors, $t(df48) = 3.64, p < .01$. The magnitude of the difference in the means of juniors and seniors in IELTS academic reading test was large (Eta squared = .21). This means that seniors could perform better and that IELTS academic reading test is a valid measure for testing academic reading with large potential of differentiation between learners.

To test the validity of VP as a measure of lexical richness against "Web-VocabProfile (BNC)" (Cobb, 2009b) as an already established measure, the correlations between VP indexes and VP (BNC) indexes of the essays were calculated. VP (BNC), another measure of lexical richness based on BNC, employed by Nation (2006) in a corpus analysis showed to be a valid measure for sorting vocabularies of a text. Therefore, VP's K1-K2 index (most frequent words in GSL) of the participants' essays was correlated with their VP (BNC)'s K1-K2 index (most frequent words in BNC corpus). Moreover, VP's AWL-OL index (less-frequent words in English, including AWL words and words not included in GSL and AWL) of the participants' essays was correlated with VP (BNC)'s K3-K20 index (less-frequent words in BNC corpus). VP's K1-K2 index for the essays correlated .64, $p < .01$, 2-tailed with VP (BNC)'s K1-K2 index, and VP's AWL-OL correlated .60, $p < .01$, 2-tailed with VP (BNC) K3-K20 index. These results suggest that VocabProfiler may be a valid measure for lexical richness and sorting out vocabularies in a text according to factors like frequency, range, register, etc. that West (1953) and Coxhead (2000) have claimed.

Testing the validity of MLTU as a measure of syntactic complexity was conducted correlating MLTU indexes of the essays with their MLS (mean length of sentence) indexes. Most of the 25 studies in Ortega's (2003) research synthesis found MLS as a valid measure of syntactic complexity. Therefore, for validation of MLTU index of the participants' essays, it was correlated with their MLS index. There was found a relatively high and significant correlation between these two measures of syntactic complexity. MLTU index correlated .62, $p < .01$, 2-tailed with MLS index. The result suggests that MLTU may be a valid measure for computing syntactic complexity of the learners' compositions.

3.3 Procedure

An academic reading test of IELTS was administered along with an opinion-essay writing task in a two-hour session. The topic of the composition was of a general nature and did not require expert knowledge of a particular subject matter. The students were instructed to produce a passage of at least 300 words in the 50 minutes time they were given. Three scores were calculated for each participant: IELTS academic reading score, MLTU, and VP indexes of the compositions. The first 300 words of each text were entered into Web-VocabProfile for lexical richness analysis. For each text, the profiler calculated the percentage of words of the text that fall into K1, K2, AWL, and OL words. AWL and OL were condensed together (AWL-OL) to show lexical richness of every essay. The essays were also analyzed syntactically using MLTU for their syntactic complexity features. The length of each essay in number of words was divided by the number of T-units in that essay, and the result was MLTU which was an index of syntactic complexity of the participants' essays and, conclusively, their abilities in producing complex syntactic structures. Finally, IELTS academic reading test scores were correlated with and regressed against MLTU and VP indexes of the essays so that the more potent predictor of academic reading would be identified.

To examine the predictive potential of VocabProfile and MLTU indexes in the regression study, IELTS academic reading scores were regressed against AWL-OL and MLTU indexes of the essays. A multiple regression analysis was employed for such purpose. That variable was supposed to be a better predictor of academic reading performance which had a higher and more significant Beta value. The level of significance adopted for this study was .05 and for all statistical calculations, version 18.0 of SPSS software was used.

4. Results

Descriptive statistics for IELTS academic reading test scores, VP and MLTU indexes of the participants' essays are presented below in Table 1. A non-parametric statistic known as one-sample, Kilmogorov-Smirnov was also run, which shows that the distribution of all scores and indexes is normal. An Asymp index in this test must be larger than .05 to show normality.

Table 1

Statistics for IELTS academic reading scores, VocabProfile & MLTU indexes in the participants essays

	IELTS	K1-K2	AWL-OL	MLTU
N	50	50	50	50
Mean	13.28*	93.93%	6.06%	13.74
SD	5.38	2.71	2.72	2.69
Asymp. Sig. (2-tailed)	.63	.59	.60	.92

Note. *Maximum Score = 40

The higher profile of K1-K2 index in the essays does not mean more lexical knowledge or richness, because this index indicates a higher profile of most frequent easier words in the texts. The small profile of AWL-OL index in the essays may be an indicator of poor lexical richness needed for academic reading, because this index indicates the percentage of less-frequent and academic words in the texts. Although there is not a particular norm for MLTU index of a text, the small figure of MLTU index in the essays may be an indicator of poor abilities for producing and comprehending complex syntactic structures because this mean is far smaller than the mean expected from advanced learners of English.

4.1 The prediction potential of VocabVrofile indexes and MLTU for IELTS academic reading performance

A standard multiple regression analysis in the Enter method was run to investigate the prediction potential of syntactic complexity and lexical richness for academic reading performance. Two independent or predictor variables were used: syntactic complexity and lexical richness. The dependent or criterion variable was academic reading performance. Preliminary assumption testing was conducted to check for multicollinearity (variance inflation factor and tolerance), normality, linearity, homoscedasticity and independence of residuals, with no serious violations noted.

Table 2

Beta Values for Regressing IELTS Academic Reading Test Scores against VocabProfile & MLTU Indexes

Model	Beta	t	Sig.
1 (Constant)		-.444	.659
AWL-OL	.386	3.141	.003
MLTU	.367	2.989	.004

Note. Dependent Variable: IELTS

The results of individual variables are displayed in Table 2. In fact, the data presented in this table tell us whether and to what extent AWL-OL and MLTU, separately, have been able to predict the variance in IELTS academic reading performance. As shown in this table, both independent variables, namely AWL-OL and MLTU, are able to predict the variance in academic reading comprehension scores. In order to interpret the results and see the weighting of the two variables that have significantly predicted the variance in the academic reading comprehension scores, we must look at the Beta (β) values. Beta indicates the effect that one standard deviation

unit change in the independent variable has on the dependent variable.

With respect to our data, $\beta = .39$, $p < .01$ for AWL-OL and $\beta = .37$, $p < .01$ for MLTU, one standard deviation unit change in the score for AWL-OL would result in .39 units of change in the IELTS academic reading comprehension scores, while with respect to MLTU, this unit of change would be .37. At first look, lexical richness appears to be a more potent predictor; however, beta values and their significance for both variables were very close, thus, we cannot say that AWL-OL is a far better predictor of academic reading comprehension than MLTU.

5. Discussion

The broad goal of our study was to discern the role that syntactic abilities, as opposed to lexical richness, plays in academic reading performance. Our main question was ‘which is the best predictor of academic reading performance: syntactic complexity or lexical richness?’ Theories and results of previous researches led us to hypothesize that syntactic complexity may be as potent a predictor of academic reading performance as lexical richness. Therefore, our focus has been on the relevance of syntactic complexity to difficulties in comprehending academic reading texts. IELTS academic reading scores were regressed against AWL-OL and MLTU indexes as independent or predictor variables. Both independent variables were able to predict the variance in academic reading comprehension scores. With respect to our data, Beta values and their significance for both variables were very close, i.e., one standard deviation unit change in the score for AWL-OL would result in .39 units of change in the IELTS academic reading comprehension scores, while for MLTU, this unit of change would be .37; thus, we could not say that AWL-OL, as a measure of lexical richness, is a far better predictor of academic reading performance than MLTU, as a measure of syntactic complexity. Moreover, we might claim that lexical richness could no longer be regarded as the single element of difficulty in academic reading texts, and complex syntactic structures make such texts difficult to comprehend even when vocabulary is familiar.

These results are to some degree consistent with the earlier findings (Cohen et al., 1979; Cobb & Morris, 2004; Martohardjono et al., 2005) suggesting that academic and general reading texts are different in their lexical and syntactic requirements; therefore, lexical and syntactic abilities can be predictors of academic reading performance. However, through further analysis, we can say that previous claims about lexical knowledge as the single predictor of reading comprehension may be doubted upon at least in academic reading performance, since the presence of complex sentences are as mentally challenging for readers in such tests as less-frequent and academic words. Findings of the present study that to some extent justify syntactic complexity rather than grammatical knowledge as significant as lexical richness in academic reading performance, may be explained by the fact that academic texts are abstract and include long, embedded, and complex syntactic structures which also need strong syntactic processing abilities instead of mere knowledge of academic words and grammatical rules. In these texts, it is not the structures themselves that make comprehension difficult, but the demands these structures make on the subsidiary processing mechanisms, especially working memory. It is a certain type of syntactic knowledge that is important for academic reading comprehension; specifically it is the ability to process complex structures such as subordinate clauses that might be directly relevant to the task.

Our study is unique in that the researcher aimed at avoiding failures of previous researches by employing measures of lexical richness and syntactic complexity. Such measures can test participants’ lexical and syntactic abilities in “context”. They test learners’ ability in producing complex syntactic structures and less-frequent words in their essays rather than merely testing their declarative knowledge of grammatical rules and vocabularies.

5.1 Implications of the study

The findings of the present research may contribute to the EAP syllabus designers, teachers, and particularly learners to deal better with linguistic features of academic texts. Based on the findings of this research,

deepening EAP learners' knowledge of AWL-OL index and improving their syntactic processing abilities may be essential factors in boosting their academic achievement. By enabling EAP learners to use more complex sentences, to embed less relevant information, to use logical connectors for integrating more sentences and to be engaged in more university studies (Ortega, 2003), MLTU scores and syntactic processing abilities of the learners may be improved substantially. Moreover, the LFP and MLTU measures can be used as diagnostic tools to identify EAP students who in their writing predominantly rely on high-frequency vocabulary and simple sentences.

5.2 Limitations and suggestions for further research

Some limitations restricted the extent of generalizing in the research. The sample was small, and the sampling technique was not random. There was also a need for a general module test of IELTS reading comprehension so that the prediction potential of VocabProfile indexes and syntactic complexity of both academic and general reading tests of IELTS could be compared. Maybe there was a need for a test of EAP reading like IELTS but with shorter texts, because of text-processing effects that IELTS reading tests have. Alderson (2000) states that longer tests allow test takers to reduce reliance on sentential processing abilities that might tap syntactic and lexical knowledge more than discourse processing abilities (p. 109). Another limitation was that IELTS academic reading tests are not representative of academic texts, since some are too specialized and some are too general (Alderson, 2000). This may be shown in the close AWL indexes of both general and academic reading tests of IELTS, i.e., 6.8 % versus 7.5 % respectively (Karami, 2012).

There are some questions that may be posed by L2 researchers and teachers. 'Would the same results be obtained in studies on other academic skills?' There is even a further question: Can we achieve the same results with nonlength-based measures of syntactic complexity, e.g., node-count, and other measures of testing vocabulary in use, e.g., Laufer and Nation's 'Vocabulary Levels Test - Productive'?

6. Conclusion

Our main goal in this study was to identify the strongest predictor of academic reading performance. The findings of this study show that both lexical richness and syntactic complexity are able to predict the variance in academic reading comprehension scores; however, these findings suggest that previous claims about lexical knowledge as the single element necessary for good performance in academic reading may be doubted upon, since processing complex syntactic structures is as mentally challenging for L2 readers as the recognition of less-frequent and academic words. The findings may be explained by the fact that academic texts are abstract, expository and include long and complex syntactic structures which also need strong syntactic processing abilities instead of mere knowledge of academic words and grammatical rules. In such texts, it is the demands that complex structures make on the subsidiary processing mechanisms, especially working memory. We may conclude that L2 readers of such texts need more the ability to process complex structures such as subordinate clauses that might be directly relevant to the task. Our study is unique in that the researcher aimed at avoiding failures of previous researches by employing measures that test the learners' ability in producing complex syntactic structures and less-frequent words in their essays rather than merely testing their declarative knowledge of grammatical rules and vocabularies. Syllabus designers, EAP teachers and readers may benefit from these findings. Deepening EAP learners' knowledge of academic and less frequent words and improving their syntactic processing abilities may be essential factors in boosting their academic achievement.

7. References

- Alderson, C. J. (2000). *Assessing reading*. Cambridge: Cambridge University Press.
<http://dx.doi.org/10.1017/CBO9780511732935>
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.

- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1(1), 1-47. <http://dx.doi.org/10.1093/applin/1.1.1>
- Carpenter, P. A., Myiake, A., & Just, M. A. (1994). Working memory constraints in comprehension: Evidence from individual differences, aphasia, and aging. In M. A. Gernsbacher (Ed.), *Handbook of Psycholinguistics* (pp. 1075-1022). San Diego, CA: Academic Press.
- Carroll, D. W. (2008). *Psychology of language* (5th ed.). Belmont, CA: Thomson Wadsworth.
- Carter, R. (2000). *Vocabulary: Applied linguistics perspectives* (2nd ed.). London: Routledge.
- Celce-Murcia, M., Dörnyei, Z., & Thurrell, S. (1995). Communicative competence: A pedagogically motivated model with content specifications. *Issues in Applied Linguistics*, 2, 5-35.
- Chapelle, P. (1994). Are c-tests valid measures for L2 vocabulary research? *Second Language Research*, 10, -157-187.
- Cobb, T. & Morris, L. (2004). Vocabulary profiles as predictors of the academic performance of Teaching English as a Second Language Trainees. *System*, 32, 75-87. <http://dx.doi.org/10.1016/j.system.2003.05.001>
- Cobb, T. (2009a). Classic Web-VocabProfile English (Version 3.0) [Computer software]. Available at <http://www.lextutor.ca/vp/eng/>
- Cobb, T. (2009b). Web-VocabProfile / BNC-20 (Version 3.2) [Computer software]. Available at <http://www.lextutor.ca/vp/eng/>
- Cohen, A., Glasman, H., Rosenbaum-Cohen, P. R., Ferrera, J., & Fine, J. (1979). Reading English for specialized purposes: Discourse analysis and the use of student informants. *TESOL Quarterly*, 13, 551-564. <http://dx.doi.org/10.2307/3586449>
- Conrad, S. M. (1999). The importance of corpus-based research for language teachers. *System*, 27, 1-18. [http://dx.doi.org/10.1016/S0346-251X\(98\)00046-3](http://dx.doi.org/10.1016/S0346-251X(98)00046-3)
- Cook, V. J., & Newson, M. (1996). *Chomsky's universal grammar: An introduction* (2nd ed.). Oxford: Blackwell Publishers.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213-238. <http://dx.doi.org/10.2307/3587951>
- Dudley-Evans, T., & St John, M. (2005). *Developments in English for specific purposes: A multi-disciplinary approach*. Cambridge: Cambridge University Press.
- Field, J. (2003). *Psycholinguistics: A resource book for students*. London: Routledge.
- Gass, S. M., & Selinker, L. (2008). *Second language acquisition: An introductory course* (3rd ed.). New York: Routledge.
- Geva, E. (1992). The role of conjunctions in L2 text comprehension. *TESOL Quarterly*, 26(4), 731-747. <http://dx.doi.org/10.2307/3586871>
- Hamp-Lyons, L. (2002). English for academic purposes. In R. Carter & D. Nunan (Ed.), *The Cambridge guide to teaching English to speakers of other languages*. Cambridge: Cambridge University Press.
- Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in Second Language Acquisition*, 21, 303-317. <http://dx.doi.org/10.1017/S0272263199002089>
- Homburg, T. J. (1984). Holistic evaluations of ESL compositions: Can it be validated objectively? *TESOL Quarterly*, 18(1), 87-108. <http://dx.doi.org/10.2307/3586337>
- Jiang, W. (2013). Measurements of development in L2 written production: The case of L2 Chinese. *Applied Linguistics*, 34(1), 1-24. <http://dx.doi.org/10.1093/applin/ams019>
- Karami, M. (2012). The relative significance of lexical richness and syntactic complexity in IELTS academic reading tests. *Iranian EFL Journal*, 8(5), 208-223.
- Kess, J. F. (1992). *Psycholinguistics: Psychology, linguistics, and the study of natural language*. Philadelphia: John Benjamin North America.
- Laufer, B. & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16(3), 307- 322. <http://dx.doi.org/10.1093/applin/16.3.307>
- Lemmouh, Z. (2008). The relationship between grades and the lexical richness of student essays. *Nordic Journal of English Studies*, 7(3), 163-180.

- Martohardjono, G., Otheguy, R., Gabriele, A., de Goeas-Malone, M., Szupica-Pyvzanowsky, M., Troseth, E., et al. (2005). The role of syntax in reading comprehension: A study of bilingual readers. In J. Cohen, K. T. McAllister, K. Rolstad, & J. McSwan (Eds.), *Proceedings of the 4th International symposium on Bilingualism* (pp. 1522-1544). Somerville, MA: Cascadilla Press.
- Mousavi, S. A. (1999). *A dictionary of language testing* (2nd ed.). Tehran: Rahnama.
- Nation, P. (1990). *Teaching and learning vocabulary*. New York, NY: Heinle & Heinle.
- Nation, P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review*, 63(1), 59-82.
- Nation, P., & Waring, R. (1997). Vocabulary size, text coverage and word lists. In N. Schmitt, & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 6-19). Cambridge: Cambridge University Press. <http://dx.doi.org/10.3138/cmlr.63.1.59>
- Nuttal, C. (1996). *Teaching reading skills in a foreign language*. Oxford: Macmillan Heinemann English Language Teaching.
- Ortega, L. (2003). Syntactic complexity measures and their relationships to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics*, 24, 492-518. <http://dx.doi.org/10.1093/applin/24.4.492>
- Paribakht, T. S., & Wesche, M. B. (1993). The relationship between reading comprehension and second language development in a comprehension based ESL program. *TESL Canada Journal*, 11, 9-29.
- Perfetti, C. A., Landi, N., & Oakhill, J. (2005). The acquisition of reading comprehension skill. In M. J. Snowling & C. Hulms (Eds.), *The science of reading: A handbook*. Oxford: Blackwell Publishers. <http://dx.doi.org/10.1002/9780470757642.ch13>
- Qian, D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning*, 52(3), 513-536. <http://dx.doi.org/10.1111/1467-9922.00193>
- Ransdell, S. (2003). The care and feeding of monolingual and bilingual university students in South Florida: Implications for assessment and training. *Psychology Learning and Teaching*, 3(2), 126-130. <http://dx.doi.org/10.2304/plat.2003.3.2.126>
- Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511732942>
- Read, J., & Chapelle, C. A. (2001). A framework for second language vocabulary assessment. *Language Testing*, 18(1), 1-32.
- Richards, J. C., & Schmidt, R. (2010). *Longman dictionary of language teaching and applied linguistics* (4th ed.). London: Longman (Pearson Education).
- Rimmer, W. (2006). Measuring grammatical complexity: The Gordian Knot. *Language Testing*, 23(4), 497-519. <http://dx.doi.org/10.1191/0265532206lt339oa>
- Scarcella, R. (2003). Academic English: A conceptual framework. *Linguistic Minority Research Institute Newsletter*. University of California, Santa Barbara.
- Shiotsu, T., & Weir, C. J. (2007). The relative significance of syntactic knowledge and vocabulary breadth in the prediction of reading comprehension performance. *Language Testing*, 24(1), 99-128. <http://dx.doi.org/10.1177/0265532207071513>
- Smith, N., & Tsimpli, L. (1995). *The mind of a savant: Language learning and modularity*. Oxford: Blackwell Publishers Ltd.
- Szmrecsányi, B. M. (2004). On operationalizing syntactic complexity. *7es Journées Internationales d'Analyse Statistique des Données Textuelles (JADT)*, 1031-1038.
- Uribe, D. (2008). Characteristics of academic English in ESL classrooms. *The Internet TESL Journal*, 14(3).
- van Gelderen, A., Shoonen, R., de Glopper, K., Hulstijn, J., Snellings, P., & Simis, A. (2003). Roles of linguistic knowledge, meta-cognitive knowledge, and processing speed in L3, L2, and L1 reading comprehension. *International Journal of Bilingualism*, 7(1), 7-25. <http://dx.doi.org/10.1177/13670069030070010201>
- West, M. (1953). *A general service list of English words*. London: Longman.
- Wilson, D. (1995). Relevance and understanding. In B. Brown, K. Malmkjær, A. Pollit, & J. Williams (Eds.), *Language and understanding* (pp. 35-58). Oxford: Oxford University Press.

Appendix A

Writing task for eliciting opinion essays

First name: Last name: Gender: male female Group: Junior Senior

You should spend about 50 minutes on this task. Write about the following topic:

“In many countries, schools have severe problems with student behavior. What do you think are the causes of this? What solutions can you suggest?”

Give reasons for your answers and include any relevant examples from your own knowledge or experience. Write more than 250 words (around 25 lines).

Appendix B

VocabProfile indexes of an essay with the highest AWL-OL profile

K1 words (1-1000):	83.76%
K2 words (1001-2000):	4.06% (<i>italicized in the following text</i>)
1k + 2k:	87.82%
AWL words:	6.09% (bold in the following text)
Off-list words:	6.09% (<u>underlined in the following text</u>)
AWL-OL:	12.18%

According to experience, many problems exist in the schools all over the world, which have many different causes. One of these problems is *behavioral* ones, which are learned in school, and friends are so important, and sometimes these *behavioral* problems are the result of bad culture or even bad movies. The second problem is family problems, negotiations between father and mother, the economic problems of the family, and using drugs in the family. These two problems should be *cured* by *referring* to an expert or consultant. Sometimes, as I mentioned, the movies and culture are the sources of these problems, so we should *correct* the scenes of the movies including action movies and blue *films* or try to reform the culture that dominates the community. In some cases, friends are so important to be considered. Because in teenagers friends can affect the personality so much, and *especially* there is a point, which is young produce, which is the other cause of these problems, and in this case, we should not use bad language and taboo or *punishment*. We should try to have friendship and affect the thoughts without any conflict. Sometimes, these problems are instinct and innate.

Appendix C**Table A.1***Exemplification of some sentences and T-units*

Example sentences	T-units	MLTU
1. I've known this woman for a long time. We can rely on her, I am confident of that. She'll get the job done.	4	6.25
2. This woman will get the job done. I've known her for a long time, so I am confident we can rely on her.	3	8
3. This man – and I've known him for a long time so I am confident that we can rely on him – will get the job done.	3	8.67
4. This man will get the job done. I am confident that we can rely on him, because I've known him for a long time.	2	12.50
5. This woman, whom I've known for a long time and whom I feel confident we can rely upon, will get the job done.	1	24