ANALYZING TOPIC-SPECIFIC VOCABULARY IN EFL STUDENT WRITING

ABSTRACT

The aim of the present paper is to report findings from a pilot study in the field of language assessment designed to investigate the criterion of lexical richness in EFL student writing. The study sets out to analyze a corpus of EFL student essays in order to investigate how different word frequency bands contribute to the topic specificity of the essays. The analysis was carried out using the Lexical Frequency Profiles proposed by Laufer and Nation. The findings of the investigation show that topic-specific words, which contribute to the lexical richness of the essays and help develop the topics, are mostly found in the bands of lower frequency. It is also suggested that the use of topic-specific words is a gradable parameter and its descriptors might be incorporated in the scoring scales used to assess learner essays.

Keywords: EFL productive vocabulary, vocabulary assessment, lexical frequency, topic-specific vocabulary

INTRODUCTION

Studies of vocabulary acquisition and assessment are definitely enjoying a revival of interest. It has long been recognized that vocabulary is one of the major underlying linguistic competences correlating well with general linguistic proficiency in a foreign language. It is the assessment and description of productive vocabulary of foreign learners of English that is at the focus of the present research study which has been designed as a pilot study to investigate the assessment of lexical range in EFL student writing with a focus on the use of topic-specific vocabulary.

Literature on L2 productive vocabulary gives several accounts of productive lexical competence. In its broadest sense, lexical competence is seen as one of the communicative language competences, more specifically, a linguistic competence, related to the learner’s ‘knowledge of, and ability to use, the vocabulary of a language’ (Council of Europe 2001: 110). It is further subdivided into receptive and productive knowledge, which is based on the difference between the ability to recognize a word when reading or listening and to use it in speech or writing (Nation 1990: 31-33). The latter division has been traditionally viewed as two extremes of the receptive-productive continuum, assuming that words move from receptive to the productive knowledge as the learner acquires more competence in a language. Yet this view causes certain argument among researchers, particularly when it comes to the definition and assessment of the underlying construct or knowledge. Hence for the purposes of vocabulary testing and assessment, Read (2000) offered to distinguish more specific components of lexical...
competence which allow looking ‘separately at recognition vs. recall and understanding vs. use’ (2000:249), of which the former is appropriate for discrete-point approach to vocabulary tests and the latter applies to context-dependent vocabulary testing and assessment. Thus, it is the ability to understand and use vocabulary of a foreign language that are evidenced and tested in EFL student writing.

The Common European Framework of Reference (CEFR) presents a taxonomy of lexical competence, which is divided into lexical and grammatical elements (Council of Europe 2001: 110-112). The latter represent the closed word classes and include articles, pronouns, auxiliaries and other high-frequency words, while the lexical elements consist of fixed expressions (sentential formulae, phrasal idioms, fixed frames and other fixed phrases, and fixed collocations) and single word forms of the open word classes (nouns, verbs, adjectives and adverbs). As a good starting point for any discussion of lexical competence, this taxonomy is designed for further specification and adaptation for specific language assessment contexts, one of which might be the assessment of lexical competence in L2 student writing.

The CEFR also provides illustrative scales for the description of the range of vocabulary knowledge and ability to control it, of which the former refers to the lexical richness and the latter to the accuracy of production. Some of these descriptors have been successfully used to design analytical assessment scales which are used to assess the writing paper of the matura examination of foreign languages in Lithuania (ŠMM & NEC 2005). The descriptors used in the scales, however, are very general and they leave much freedom of interpretation, which might have a negative effect on the reliability of assessment. Experience in using these assessment scales proves that no matter how much attention is paid to the application and interpretation of the scales during the pre-examination training of assessors, every assessment session reveals the necessity to specify them in greater detail. It is therefore important to establish more explicit parameters of lexical range which would be easy and simple enough to use and design in different writing assessment situations.

One of the possibilities to derive such parameters is to analyze the components of the process of writing and its product, i.e. written texts. Weigle (2002) provides a thorough discussion of different models of writing, of which Grabe and Kaplan’s model is singled out as offering the most rigorous taxonomy of academic writing skills, knowledge and processes (Grabe & Kaplan 1996: 217-222). In their model, Grabe and Kaplan approach the task of writing in terms of participants, setting, task, text and divide language knowledge into linguistic, discourse and sociolinguistic knowledge. Being a component of linguistic knowledge, vocabulary knowledge consists of (1) interpersonal words and phrases, (2) academic and pedagogical words and phrases, (3) formal and technical words and phrases, (4) topic-specific words and phrases, and (5) non-literal and metaphorical language (ibid 1996: 220). Hence, it is suggested here that an analysis of L2 written production in terms of these components might yield reliable parameters for the assessment of lexical range.

One of the most influential studies of lexical richness of L2 learner writing was done by Laufer and Nation (1995). They proposed a new measure of lexical richness, the Lexical Frequency Profile, which is based on the parameter of word frequency. This tool analyzes a piece of writing in terms of word frequencies and produces word frequency profiles which give an account of proportions of words across different frequency levels. Despite certain critique (Meara 2005), the tool has won recognition and proved to be a reliable measure of lexical use in writing (Laufer 2005).

The data generated by the Lexical Frequency Profile could be also analyzed to define yet another parameter of lexical richness, i.e. topic specificity. It is argued here that the word
frequency analysis could be related to the expression of the topic of a learner essay and inform about his/her ability to produce words related to one or another topic or semantic field. This parameter of lexical richness or vocabulary shall be termed topic specificity of vocabulary. It is also hypothesized that topic-specific vocabulary belongs to the words of lower frequency and thus allows making inferences about the quality or depth of vocabulary competence.

Vocabulary competence or the lexical richness of a student text is often described in rather fuzzy terms, such as ‘sophisticated’, ‘adequate’ or ‘limited’ range. Possibly, if a text is written on a given topic, then the ‘adequacy’ might be defined in terms of topic-specific words and respective descriptors could be formulated. Therefore an analysis of the use of topic-specific vocabulary might shed light on the learner’s ability to produce words on a given topic and highlight certain parameters which might be taken into account when making any evaluative judgments about the lexical richness of a student essay.

GOALS OF THE WORK

This research study has been designed to analyze the vocabulary of EFL learner essays and it aimed at the following goals:

1. To establish which level of lexical frequency contributes most significantly to the topic specificity of EFL student essays.
2. To compare how the use topic-specific vocabulary varies across different frequency bands and different topics.
3. To compare the use of topic-specific vocabulary in low- and high-rated essays.

MATERIALS AND METHODS

To achieve the goals of the study, a corpus of thirty student essays (ca. 11,900 words) was analyzed. The essays were written by first-year students of the Department of English Philology of Vilnius University as in-class examination essays in the period 2004-2006. They were marked by two teachers and holistic scores ranging from ‘5’ as the lowest mark to ‘10’ as the highest were given.

To make a comparison of the use of topic-specific vocabulary possible, the corpus was compiled of essays on three topics: Topic 1 ‘Capital punishment should/should not be introduced in Lithuania’; Topic 2 ‘Sex education should/should not be introduced in secondary schools’; Topic 3 ‘Several generations should/should not live under one roof’. The topics of the examination essays were taken from the course syllabus and therefore were familiar to the students who were expected to have enough vocabulary to produce adequate argumentative essays.

The handwritten examination papers were typed to make them available in the electronic format and analyzed using the Web Vocabulary Profiler (version 2.5 Classic) which is freely available on the internet website at http://132.208.224.131. The Vocabulary Profiler is a tool designed by T. Cobb and based on the Lexical Frequency Profiles (Laufer and Nation 1995). It breaks down any pasted text into four frequency bands where:

‘K1’ is the first thousand of the most frequent English words;
‘K2’ is the second thousand of the most frequent English words;
‘AWL’ is the Academic Word List which includes the most frequent vocabulary of academic texts across different disciplines;
‘Off-list words’ include all the remaining words which do not fall within any of the previous bands and have lower frequency in the language.

Apart from the four sub-lists, the output page of the Vocabulary Profiler also provides statistical data which among other parameters include the number of running words or tokens, the number of different words or types, type-token ratio, lexical density etc.

For the purposes of the study, lexical frequency profiles were computed for each individual essay and for the three subcorpora of topic-specific essays and the output was analyzed.

It should be remarked that when typewriting the essays, spelling errors where treated. It is assumed here that the assessment of a learner’s ability to understand and reproduce a word in writing should be separated from the assessment of the accuracy of reproduction. Besides, this is often accommodated by writing assessment scales, where lexical range is assessed separately from language accuracy. Hence, a student may be penalized for a spelling error and be given points for the reproduction of a word. Therefore the analysis allows making inferences about a learner’s ability to use words in writing irrespectively of the accuracy of spelling. Words, which were derived using wrong affixes, however, demonstrate that a learner is not able to form correct derivative forms and therefore were excluded from frequency counts. It refers, for instance, to the following items: Topic 1 essays: *barbaristic (instead of barbarous or barbaric), *murdery (‘murder’), *raper (rapist), *unhuman or *inhumanous (inhuman); Topic 2 essays: *contraceptics (contraceptives), *precautional (precautionary); Topic 3 essays: *contra-arguments (counterarguments).

RESULTS

1. General tendencies

The analysis of the lexical frequency profiles showed that essays on different topics were similar in the majority of parameters (Table 1).

Table 1. Lexical frequency parameters of different topic subcorpora

<table>
<thead>
<tr>
<th></th>
<th>Topic 1</th>
<th>Topic 2</th>
<th>Topic 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of words</td>
<td>4049</td>
<td>3684</td>
<td>4160</td>
</tr>
<tr>
<td>Number of types</td>
<td>822</td>
<td>727</td>
<td>853</td>
</tr>
<tr>
<td>Type-token ratio</td>
<td>0.20</td>
<td>0.20</td>
<td>0.21</td>
</tr>
<tr>
<td>Tokens per type</td>
<td>4.93</td>
<td>5.07</td>
<td>4.88</td>
</tr>
<tr>
<td>Lexical density</td>
<td>0.51</td>
<td>0.54</td>
<td>0.55</td>
</tr>
</tbody>
</table>

As seen from Table 1, Topic 2 had a somewhat smaller number of different types, which might be related to the comparatively shorter length of essays on this topic with an average length of 368 words as compared to 404 in Topic 1 and 416 in Topic 3.

The three subcorpora, however, differed in the distribution of words across the four frequency bands (Table 2).

Table 2. Distribution of vocabulary across different frequency bands in different topic subcorpora (in percentages)
The analysis showed that essays written on Topic 3 contained the smallest number (1.92 per cent) of Off-list words, i.e. words belonging to the bands of the lower frequency. It may be therefore assumed that certain topics are less demanding in terms of specific vocabulary and they do not stimulate students to resort to more complex lexical expression. The proportion of K1 words in this subcorpus (85.55 per cent) also indicates that the use of the topic-specific vocabulary was compensated by high-frequency lexis.

2. Topic-specific vocabulary across different frequency bands

The analysis showed that topic-specific vocabulary is best evidenced in the bands of lower frequency. It is the Off-list words that include the largest amount of topic-specific vocabulary. Moreover, these words are members of semantically related word groups. Table 3 illustrates the Off-list items in the three topic subcorpora.

Table 3. Low frequency (Off-list) words used in the student essays

<table>
<thead>
<tr>
<th>TOPIC 1 (113 types; 289 tokens)</th>
<th>K1</th>
<th>K2</th>
<th>AWL</th>
<th>Off-list</th>
</tr>
</thead>
<tbody>
<tr>
<td>abolish (-ment), abstain, absurd, accomplish (-ment), alcohol, America, amnesty, assassination, assert, ban, Bible, bishop, budget, burden, cancel, career, Catholic (-ism), chamber, Christian (-ity), cigarettes, convict (n/v; -ed), corruption, counterargument, criticize, cruel, democracy (democratic), deprive, destiny, devious, disastrous, discourage, etc, execute, fame, fulfill, gas, grief, holy, huge, humanity, immoral, imprison (-ed, -ment), inefficiency, inflict, inhuman, injection, innocent (-ence), jail, Jesus, juvenile, killing, legalize (-ation), legitimate, Lithuania (-n), maniac, manslaughter, Muslim, offence, offender, optimistically, outweigh, penalty, pension (-er), plague, pope, practice, proponent, regain, scare, segment, sentenced, serial, sinful, stimulate, terrify, testament, unsatisfactory, vast, vicious, victim, weighty, yoga.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TOPIC 2 (75 types; 207 tokens)</th>
<th>K1</th>
<th>K2</th>
<th>AWL</th>
<th>Off-list</th>
</tr>
</thead>
<tbody>
<tr>
<td>abortion, absurd, acquainted, adolescence, ajurveda, American, budget, Canada, carrier, casual, censured, competitive, condoms, contraception, cope, curricula, delinquency, drugs, embarrassed, enlighten (-ed), essay, etc, faintest, gonorrhea, horrible, infect (-ed, -ions), inspire, intercourse, intimate, juvenile, kindergarten, lifestyle, magazine, masturbation, meditation, menacing, organism, oriental, penis, pills, practice, precaution (-ary), pregnant (-ancy), puberty, redundant, refuted, sentencing, shy, syphilis, taboo, teenage (-er), teens, testicles, TV, UK, underage, unexpected, unintended, USA, vagina, venereal, victims, vital, worldwide, yoga.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TOPIC 3 (67 types; 80 tokens)</th>
<th>K1</th>
<th>K2</th>
<th>AWL</th>
<th>Off-list</th>
</tr>
</thead>
<tbody>
<tr>
<td>abuse, addicted, adolescent, advantageous, alien (-s, -ated), anecdotal, apartment, atmosphere, beloved, budget, calmer, celebrations, cell, champagne, chaos, comprehend, compromise, convey, cope, cozy, disabled, disco, disrupt, divorce, emergency, enduring.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
essay, etc, exaggerated, exhausting, expenditure, friction, household, impartial, intimate, irritate (-ing), misunderstanding, mother-in-law, opposition, overstep, pathological, pension, personality, privacy, provoke, rituals, routine, scared, segment, socialize, sorrows, spouse, surroundings, teenagers, trivia, unavoidably, understandable, upbring, verbal

The lists of the different topic subcorpora demonstrate the prevalence of semantically related words which allow making rather accurate guesses of the topic at hand. For example, in the Topic 1 subcorpus, words like abolish, assassination, convict, cruel, imprison (-ed, -ment, legalize, manslaughter, offence, penalty, victims etc. clearly belong to one semantic group of words related to crime and punishment. Another groups of words is related to broader moral and religious issues of death penalty, e. g. Bible, burden, Catholic (-ism), criticize, democracy, humanity, immoral, inhuman etc. Admittedly, the borderline between a topic-specific and topic-neutral word may be rather arbitrary. In order to establish, whether a word contributes to the topic specificity of the text, the contexts and collocations in which a word is used in the essay were analyzed. In this way it was found that, for instance, the words chamber and gas (Topic 1) were used by the author in the collocation gas chamber, which was mentioned in the essay as a form of execution. Hence, both words might be considered to be contributing to the topic specificity of the text. An analysis of the Off-list bands of the three subcorpora showed that these bands are dominated by topic-specific items whereas the high frequency bands contain many more general and topic-neutral words and a much smaller proportion of topic specific vocabulary.

Even though both K1 and K2 lists on the same topic include lexical words that have direct associations with the topic of the essay, such words are outnumbered by functional or topic-neutral lexical words. Of all the words of the K1 band of Topic 1, the following might be considered to be topic-specific: capital, court, death, kill (-er), prevent, and right (-s). They account for only 4.64 per cent of all running words or tokens. The K2 words include a larger proportion of topic-specific vocabulary, e.g. accused, crime(s), guilty, mercy, murder (-er(s)), prison (-er), punish (-ment) and account for 41.77 per cent of all the tokens in this band. In comparison, in the Off-list band of Topic 1, 87.5 per cent of words might be considered to be topic-specific. Thus, it is obvious that the prevalence of topic-specific vocabulary tends to increase in the bands of lower frequency. This tendency was also observed in the Topic 2 and Topic 3 subcorpora, where topic-specific words were even scarcer than in the bands K1 and K2.

As regards the AWL band, which falls between the first two thousand of most frequent English words and the Off-list words and which is a specialized list of academic language, it was not found to be contributing to topic-specificity of the essays. The list itself consists of 570 word families which are frequently used in academic texts of different disciplines (Coxhead 2000; Read 2000). Generally speaking, this layer of students’ vocabulary is undoubtedly indicative of their linguistic proficiency yet its occurrence in L2 written texts tells more about a learner’s ability to create a formal academic argumentative text than allows making inferences about his/her range of lexical expression on a given topic. Unsurprisingly, the essays of the three topic subcorpora had similar words of the Academic word list with the following words or their derivatives represented in all the subcorpora: adult, aspect, conclusion, consequence, controversy, create, culture, debate, despite, evidence, issue, maintain, sum, topic and traditional.

As it has been mentioned before, Topic 3 yielded the lowest proportion of low frequency words. Possibly, this topic was not very difficult in terms of special lexis and therefore it did not ‘push’ the students towards the upper reaches of their lexical competence. Instead, the students found it rather easy to discuss the problems of different generations living together in fairly
simple high-frequency words. Hence, Topic 3 essays were found to have the largest proportion of K1 words (85.55 per cent, see Table 3) and were the longest of the three topic subcorpora. The fact that a writing test does not impose any lexical constraints might be quite acceptable if the focus of the test is, for example, on text organization while language use and lexical range are given less importance. As it will be shown later, an attempt to relate essay scores with the use of topic-specific vocabulary proved that the scores of the essays do not correlate with the occurrence of low-frequency vocabulary.

3. The use of topic-specific vocabulary in low- and high-rated essays

The analysis showed that, in general, essays that scored higher marks contained a bigger proportion of the Off-list words than those rated lower. Table 4 presents ranges of the Off-list vocabulary proportions for the three subcorpora and quotes proportions of such vocabulary in individual high-rated essays. It should be also noted that the essays were scored holistically and the scoring was mainly focused on text organization and the development of the topic rather than certain linguistic features of the text.

Table 4. Proportions of Off-list words in high-rated essays

<table>
<thead>
<tr>
<th>Subcorpus</th>
<th>Range of the Off-list vocabulary proportions in percentages</th>
<th>Percentages of Off-list words in essays scored '8', '9' and '10'</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topic 1 subcorpus</td>
<td>4.14 - 10.08</td>
<td>6.50; 8.33; 9.36</td>
</tr>
<tr>
<td>Topic 2 subcorpus</td>
<td>2.87 - 9.44</td>
<td>4.73; 6.68</td>
</tr>
<tr>
<td>Topic 3 subcorpus</td>
<td>1.15 - 2.34</td>
<td>1.15; 1.59; 1.87; 2.05; 2.14</td>
</tr>
</tbody>
</table>

In general, essays which were given high marks contained a comparatively bigger number of low-frequency words. Yet several exceptional cases were found and since such cases might be more prominent in a larger corpus of essays, they deserve a special note. Such exceptional cases refer to essays which contain a very low proportion of low-frequency vocabulary (e.g. 1.15 per cent in Topic 3 subcorpus) and which were given high marks or, vice versa, essays having a large number of topic-specific words and given lower marks. The analysis of such cases showed that holistic scores do not correlate with the statistical parameters of lexical frequency.

It was found that each subcorpus had one or two essays which were given low and very low marks (‘5’ and ‘6’) and which had quite many words of lower frequency. The authors of those essays demonstrated little knowledge of text organization and did not develop their paragraphs properly, which was penalized when marking, but evidently they attempted to compensate their failing writing skills by giving a selection of topic-specific words. It was also observed that the low-rated essays are much shorter in length than the average. Consequently, the proportion of low-frequency vocabulary in such essays turned out to be very high. Hence, even though the use of low-frequency vocabulary indicates a higher level of lexical competence, it does not necessarily correlate well with a learner’s ability to produce written texts.

These findings, in fact, confirm the disadvantages of holistically derived scores reported in language testing literature. Weigle (2002) and Weir (2005) agree that holistic scoring is particularly difficult to use in second-language testing situations because it does not take into
account the fact that different aspects of linguistic competence develop at different rates. Hence, a learner may have poor writing skills and a high level of lexical competence. In contrast, analytic assessment scales require looking at different criteria and giving individual marks on each of them. Therefore the use of topic-specific vocabulary would be best taken into account in analytically derived essay scores.

Further research, however, is needed to see how the descriptors of the use of topic-specific vocabulary could be formulated and incorporated in the scoring scales. Possibly, lexical frequency data could be used to compile a reference list of topic-specific vocabulary to be used for the assessment of test papers, which involve production of written texts on a given topic, or attached to the scoring scales for the use of the raters. Depending on the testing context, such a list might be compiled on the basis of course material or, in case of a high stake test, after trialing the writing task. In this way, explicitly spelled-out expectations of the testers might contribute of the objectivity of assessment.

4. Limitations and implications of the study

One of the major limitations of the current study is related to the interpretation of frequency data for the purposes of assessment. The output of the Vocabulary Profiler disregards or does not give any information about a learner’s ability to reproduce multiple word units, i.e. fixed expressions and collocations, which are undoubtedly indicative of a more advanced level of lexical competence. The occurrence of a single word needs to be related back to the text to establish its context and then make inferences about the level of lexical competence. It has been found that a topic-specific word is often used in a phrase or collocation consisting of words of different frequency, e.g. death penalty, where death belongs to the band of the most frequent words and penalty is a less frequent word. Therefore, it is assumed that lexical frequency data should be seen as primary material which requires further adaptation to be used for assessment purposes.

The study was also limited in terms of corpus length and variety of topics. Therefore, a more extensive research project on a larger corpus of L2 written production is needed to collect more empirical material to confirm the findings across a larger variety of topics.

Further research is also needed to see how the descriptors of the use of topic-specific vocabulary could be incorporated in the rating scales and if a reference list of topic-specific vocabulary could be a useful tool in the scoring procedure.

CONCLUSIONS

The analysis of vocabulary of L2 student writing in terms of topic-specific words showed that topic-specific words are found in the all the four frequency bands of the Vocabulary Profiler, which is based on the Lexical Frequency Profiles, a measure of lexical richness proposed by Laufer and Nation. It is, however, the band of low-frequency words that includes the largest proportion of such vocabulary and is considered to be contributing most significantly to the topic specificity of the essays.

The fact that certain topic-specific words are found in the first two thousand of most frequent English words, which are acquired in the first stages of learning English as a foreign language, suggests that the parameter of topic specificity might be evidenced at different achievement levels and is gradable. Thus, it may be worthwhile undertaking a more elaborate
research study in order to propose descriptors of the use of topic-specific vocabulary which could be included in the analytic scoring scales for various testing and assessment contexts.

Another conclusion is related to the selection of topics for writing tasks. The analysis showed that different topics are not equally demanding in terms of lexical expression. Therefore, if a tester is interested in measuring the level of lexical range, special attention should be given to the formulation of the topic so that it would encourage the test-taker to resort to the depth of his/her lexical competence.

And finally, the research findings showed that the use of low-frequency vocabulary, which, in fact, indicates a higher level of linguistic proficiency in a foreign language, does not correlate well with holistically given marks. Thus if a writing test is meant to be a reliable measure of EFL written production, the analytic scoring scales are to be used, which among other criteria, should include the criterion of lexical range defined.

REFERENCES


Information about the author:
Lect. Rita Juknevičienė
Department of English Philology
Faculty of Philology
Vilnius University
Universiteto St. 5
Vilnius Lithuania
Tel./Fax +370 5 2687 22
rita.jukneviene@takas.lt