Effects of Lexical Density and Lexical Variety in Language Performance and Proficiency

Bonala Kondal, Research Scholar, School of ELE, EFL University, Hyderabad

ABSTRACT

The aim of the present study is to explore the effects of lexical density and lexical variety in language performance and proficiency of the written scripts of tenth class regional medium students. The exam scripts are collected as data, the learners aged between 14 and 16 years. The main purpose of the study is to establish the use of lexicon in the written scripts would exhibit language knowledge of the learners. The cognitive theory of learning used in the present study to ascertain the use of measures i.e. lexical density and variety would aid in the language learning. Brown’s (1973) observations showed that there was a predictable order of acquisition of certain morpheme in English. Jean Ure’s (1971) basic method of lexical density is used to analyze the data. Robinson’s (1995) measure of complexity is used to find out the lexical variety. This reliable and valid measure of lexical density and variety in the written scripts will be useful for determining the factors that affect the language performance and proficiency which is related lexicon.

Keywords
Lexical density, Lexical variety, Language performance, Language proficiency, Vocabulary, Lexicon.

1. INTRODUCTION

The purpose of this paper is to explore the effects of lexical density and lexical variety in language performance and proficiency through written scripts produced by regional medium students during their study. Lexical development is explored using measures such as lexical variation (vocabulary range), lexical density (proportion of content words) (Viadkovic & Barker, 2009).

The two measures that are often used to describe lexical development particularly are lexical density and lexical variety. Lexical density is a measure to find out how many lexical items such as nouns, verbs, adjectives, and adverbs used in the text, while lexical variety provides a measure of how many different words used in a text. Both the lexical density and lexical variety have been shown significantly higher in writing than in speaking (Halliday, 1989). The researcher used these two measures to explain and establish that these measures largely affect the language performance and proficiency of the learners.

Although the second language acquisition process is understood better by comprehending the human brain processes and how it learns new information, mainly focusing on the learner as an individual. As the lexicon is a part of it and it is chosen based on the assumption that learning component of second language concentrates on transition theories to property theories of the framework. The focus here is largely as an individual learner (Mitchell & Myles, 2004, p. 95). In fact, the lexicon may be the most prominent language component for learners in the development of the language. The processing component of language learning includes within the frameworks. However, Mitchell & Myles (2004) state that ‘the learner is seen as operating a complex processing system that deals with linguistic information in similar ways to other kinds of information’ (p.97). Cognitive theory recognizes learning strategies as one of the prominent cognitive processes in L2 acquisition.

The cognitive theory of learning is largely based on the theory of human information processing which deals with mental processes involved in learning. This theory mainly refers to three basic cognitive components of learning: how knowledge is developed, how knowledge becomes automatic and how new knowledge is integrated into an existing cognitive system of the learner (Takac, 2008, p.27). Prominence is given to meaningful learning, i.e. learning with comprehension which is not marked in behavior. This theory defines L2 acquisition as a complex cognitive skill which like other such skills engages cognitive systems such as perception, memory and information processing to overcome limitations in human mental capacity which may inhibit performance (Ellis, 2000, p. 175). The theory also suggests that linguistic codes and structures are stored and retrieved from the memory precisely in the similar way as other types of information.

This paper attempts to find out the role of lexical content in the language performance. In addition to that it also finds out that how the lexical density leads to proficiency of the learners and how lexical variety does affect the performance of the learners.

The study attempted to discover answers to the following research questions:
i. What is the role of lexical content in language performance among regional medium 10th class students?
ii. Does lexical density lead to proficiency of the learners?
iii. Does lexical variety affects the performance of the learners?

2. LITERATURE REVIEW

This section provides the past works used previously and also presents related theoretical framework. Nevertheless, the researchers argue that learners acquiring language go through various developmental stages. With the help of these stages learners enhance their language performance through lexical density and lexical variety. These stages are similar across the languages.

The small scale investigations in the review would provide the difference between the written work of native speakers and foreign students in terms of lexical density and lexical variation. The study used a basic method to find the lexical density in the analysis of written work in English produced by Swedish students. Though, it was done in order to apply certain theoretical methods of evaluation to essays and to use the results. In this study on the both counts the native speakers had high results. The learners’ language was richer in lexis and more variety (Linnarud, 1973).

Johansson (2008) suggested that lexical variety is a better measure to use of identifying dissimilarities between age groups. Two measures were compared to describe the lexical development. And it also important to be shown that both lexical density and lexical variety can be used for modality dissimilarities and development dissimilarities. The researcher also observed that both measures were used on the same material reveals that they were not interchangeable.

In another report the lexical frequency profile shown to be a reliable and valid measure of lexical use in writing. This profile provides similar stable results for two write ups produced by the same person, and discriminates between different learners’ proficiency levels, and it can be reasonably expected vocabulary size based on their production of the language. It also correlates well with an independent measure of vocabulary knowledge (Laufer & Nation, 1995).

3. METHODOLOGY

3.1 Participants
Ten tenth class regional medium students’ exam scripts were collected from a teacher for the study. Their age is between 14 and 16 years and they had initially five years of English exposure. The subjects’ mother tongue is Telugu. The sample consists of three female and seven male. The nature of data was a spontaneous individual production; hence the sample of data was naturally-occurred. The subjects were blend of good, average and poor at their writing abilities.

The collected students’ answer scripts were written during their pre-final examination. The most of students attempted all the questions. Then, the data was calculated to know the number of lexicon used in the scripts including both functional and content/lexical words. The length of the texts ranged between 136 and 350 words. Afterwards, all the lexical items used in the scripts were calculated.

3.2 Procedure
Jean Ure (1971) used the term, lexical density and basic method to find out the lexical density of the scripts. In addition to that this method used to describe the percentage of lexical words. All the words are calculated and a single word in the script is treated as an orthographic word (OW). Afterwards, all the content words are calculated in spite of their repeated occurrences in the text. The total number of orthographic words and the total number of lexical words (LW) are arranged in a relation to each other to accomplish lexical density:
\[
\text{Lexical Density} = \frac{\text{LW}}{100} \times \frac{\text{OW}}{\text{LW}}
\]

This formula is derived from Robinson’s (1995) study. (See table 1 measure of complexity).

<table>
<thead>
<tr>
<th>S. no</th>
<th>Measure</th>
<th>Definition</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Lexical – Type-token ratio</td>
<td>The total number of different words used (types) divided by the total number of words in the text (tokens).</td>
<td>Robinson(1995)</td>
</tr>
</tbody>
</table>

Note: This measure of complexity is adopted from Analyzing learner language by Ellis & Barkhuizen (2005).

Lexical variety count showed the percentage of the total number of content words which were used in the text notably nouns, verbs, adjectives and adverbs which were provided in table 2. Actually, these lengths vary from one lexical item to another.

The number of content words (CW) namely nouns, verbs, adjectives and adverbs. Then the total number of orthographic words (OW) are put into relation to each other to accomplish lexical variety, similarly, separate
counts are carried out to find out the percentage of nouns, verbs, adjectives and adverbs in the text.

\[
\text{Lexical Variety} = \frac{\text{CW} \times \frac{1}{10}}{\text{OW}} = \frac{100 \times \frac{1}{10}}{1} = 100
\]

With respect to lexical variety, the mean and standard deviations were calculated individually for all the content words (refer table 2 in results section).

4. RESULTS

The results for the lexical density in the study are as follows:

The length of text ranged between 136 – 350 words per script. Two of ten scripts lie below 200 words, in contrary to that only two subjects produced more than 300 words. The range of the content words are from 46 -172 words. 50% of the participants were used above 100 content words in their text, remaining produced less than 100 words. The minimum lexical density is 33.82% to the maximum of 49.76%. The study shows that eighty percent of students’ lexical density lies above 40% and remaining twenty percent of the students’ lexical density lie between 32% - 39%. In the figure 1 lexical density is indicated in the bottom line which is almost stable throughout all the scripts.

The lexical variety count provides a valuable check on the use of various content words in the text to be acquainted with the learner. In some cases the language may seem poor. Actually, repetition of nouns and verbs are more common in the written texts. Moreover, for all the nouns, verbs, adjectives and adverbs mean and standard deviations were calculated (refer table 2).

<table>
<thead>
<tr>
<th>S. no</th>
<th>No of nouns</th>
<th>No. of verbs</th>
<th>No. of adjectives</th>
<th>No. of adverbs</th>
<th>Total no of content words</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>42</td>
<td>19</td>
<td>7</td>
<td>5</td>
<td>73</td>
</tr>
<tr>
<td>2</td>
<td>78</td>
<td>42</td>
<td>23</td>
<td>4</td>
<td>147</td>
</tr>
<tr>
<td>3</td>
<td>48</td>
<td>20</td>
<td>9</td>
<td>1</td>
<td>78</td>
</tr>
<tr>
<td>4</td>
<td>54</td>
<td>21</td>
<td>13</td>
<td>4</td>
<td>92</td>
</tr>
<tr>
<td>5</td>
<td>46</td>
<td>17</td>
<td>12</td>
<td>2</td>
<td>77</td>
</tr>
<tr>
<td>6</td>
<td>42</td>
<td>25</td>
<td>9</td>
<td>2</td>
<td>78</td>
</tr>
<tr>
<td>7</td>
<td>38</td>
<td>28</td>
<td>10</td>
<td>3</td>
<td>79</td>
</tr>
<tr>
<td>8</td>
<td>60</td>
<td>19</td>
<td>11</td>
<td>1</td>
<td>91</td>
</tr>
<tr>
<td>9</td>
<td>43</td>
<td>26</td>
<td>5</td>
<td>3</td>
<td>77</td>
</tr>
<tr>
<td>10</td>
<td>24</td>
<td>14</td>
<td>3</td>
<td>1</td>
<td>42</td>
</tr>
</tbody>
</table>

Mean: \[
\frac{475}{10} = 47.5 \quad \frac{231}{10} = 23.1 \quad \frac{102}{10} = 10.2 \quad \frac{26}{10} = 2.6 \quad \frac{834}{10} = 83.4
\]

SD: 4.3, 2.36, 1.63, 0.42

The results of the lexical variety of the study are as follows:

![Figure 1](image1.png)

**Figure 1.** Lexical density is shown in terms of total no. of words and content words.

The results of the lexical variety of the study are as follows:

![Figure 2](image2.png)

**Figure 2.** Lexical variety of nouns, verbs, adjectives and adverbs are used in the text.
5. DISCUSSION

Brown’s (1973) observations showed that there was a predictable order of acquisition of certain inflectional morphemes in English. Though, the fact that they did not always occur at exactly the same age, but his study on Adam, Sarah and Eve learned English morphemes in the similar passion. Even in this study also showed the fact that if the learners have more exposure to the language they become more competent and proficiency in that language. The learners initially, used more nouns and verbs than adjectives and adverbs because it is easy to remember and generate when it is essential.

Nouns are categorized into two types like concrete nouns and abstract nouns while calculating nouns in the analysis section. Similarly, verbs divided into action verbs and dynamic or/stative verbs (see appendix). Because of inadequate number of adjectives and adverbs used in the text made the researcher to difficult for categorization. Brown (1991) measured complexity in terms of the extent to which learners hypothesized (Ellis & Barkhuizen, 2005, p. 154).

It is clearly shown in the scripts that one of learners used many number of nouns, verbs and adjectives and a few adverbs (see appendix). Almost all the learners more often used many nouns and verbs in their scripts. Further 50% of the learners did not produce many adjectives in their texts. Conversely almost all the students are unsuccessful to employ adverbs in their writing. It also elucidate that the learners possess more knowledge about nouns and verbs than other lexical items including adjectives and adverbs.

The mean scores of the lexical variety of nouns, verbs, adjectives and adverbs are 47.5, 23.1, 10.2 and 2.6. The standard deviations for the nouns, verbs, adjective and adverbs are 4.3, 2.36, 1.63 and 0.42. At this point the larger standard deviations indicate greater variability in the data, simultaneously the smaller standard deviations indicate less variability in the data. It is found that the results that standard deviation values of nouns, verbs, adjectives and adverbs are declining in the identical manner. Hence, the variation of the lexicon also arrives down according to the standard deviation.

These results show that less proficient learners make more repetitions of most frequently used words i.e. nouns and verbs. There is a tendency for less proficient learners to make use of the vocabulary. These differences are in accordance with the concept of language proficiency that assumes that richer lexical density and high lexical variety are characteristic of better language knowledge (see table 2 and figure 1).

As the scripts are of different lengths, the results of lexical variety for each are incomparable. In an elimination of this component the scripts were divided according to the number of lexical words namely high, average and low (see appendix). More number of nouns and verbs used in these three scripts (see figure.3).

6. CONCLUSION

The paper has shown that both the lexical density and lexical variety can be played an effective role in the language performance among the regional medium tenth class students. Although, the lexical density lead to the proficiency of the learners and lexical variety affects the performance of the learners as a whole. The students can be identified in terms of lexical density and variety. If their language is richer in lexis and more varied, which suggests that more emphasis needs to be given to the study of vocabulary in teaching of English. Generally, learners have a wide range of receptive vocabulary than productive vocabulary.

The study also presents that it can be expected learners’ lexical development as measured by lexical density and lexical variety to be reflected in the learner’s productive use of language performance and proficiency in the use of lexicon. The learners have to make use of language knowledge and regulate it accordingly.

It is evident that these two lexis counts do not cover the whole field of lexis. A study based on other ways such as frequency of lexis in the text and to determine the occurrences of words in the text and their difficulty level would be interested for further area of exploration.
REFERENCES


