Lexical Inferencing: The Relationship between Number and Density of Lexical Items and L2 Learners' Reading Comprehension Achievement

Nastaran Chegeni
English Department, Shahreza Branch, Islamic Azad University, Iran

Omid Tabatabaei
English Department, Najafabad Branch, Islamic Azad University, Iran

Abstract—Lexical inferencing denoted a process of guessing the meaning of an unknown word by employing all linguistic cues available in the text together with the reader’s world knowledge, his/her linguistic knowledge, and his/her awareness of the context. (Haastrup, 1991). This study dealt with exploring the influence of number and density of unknown words on lexical inferencing. To this end, different data collection devices were used: A Quick placement test given to 90 students to select 30 same English proficiency level students for this study, a text with high density of unknown words including 357 words totally and 20 unknown words and a text with low density of unknown words including 291 words totally and 10 new words were given to the participants to show the effect of number of new words on success of learners’ lexical inferencing. The guessing success of the students in the texts with low density of unknown words interpreted that the lower the number of unfamiliar words, the higher the available clues for the participants to use for inferring the correct meaning of those words. The results can be of significance to teachers and learners of English.

Index Terms—lexical inferencing, number of lexical items, reading comprehension

I. INTRODUCTION

One of the essential factors that affect language learning in general and reading in particular, is vocabulary size. According to Laufer (1989), although a foreign/second language reader cannot master the vocabulary the same as a native speaker, it is not correct to say that he/she is not able to understand the written materials.

Hu and Nation (2000) maintain that it is very good if EFL/ESL readers can cover 98 percent of the reading comprehension without help. In the same way, Hirsch and Nation (1992) suggest that “the 98% coverage can help the reader to comprehend the written text for pleasure without any assistance.” Of course, it is not easy to set the best criterion for the EFL/ESL reader although there are two criteria for the amount of known vocabulary. However, it is assumed if EFL/ESL reader can cover at least 95% of the words, it means that they can understand most of the written text; that is, they need to have the knowledge of 3000 word families, or 5000 lexical items. However, it is better to say that “[the] higher the coverage, the lower the density of unknown words” (Laufer, 1997). Considering the great importance given to lexical inferencing in many studies (Read, 2000), the main purpose of this paper was to investigate the influence of number of unknown words on lexical inferencing.

II. REVIEW OF LITERATURE

Reading Comprehension

Souvignier & Moklesgerami (2006) define reading comprehension as the reader’s ability “to read and remember, reproduce, learn from, and find deeper meaning in text for later use.” Moreover, in the process of reading the reader not only needs to comprehend the direct meaning of what he/she is reading, but, he/she also needs to understand the implied meaning of the text. According to Tierney and Readence (2005), “Learning to read is not [only] learning to recognize words; it is [also] learning to make sense of texts”. (Karbalaei, 2010, P.166). Pressley (2002a) states that reading involves a lot of cognitive capacity which is available for understanding the reading materials.

Some researchers believe that readers are required to employ a posteriori knowledge in examining the text and form new ideas while reading for comprehension. According to Guterman (2003), “the more knowledge a person brings to his or her reading, the more he or she will understand the text.” Some other researchers (e.g., Lau & Chan, 2003) maintain that for a successful reading comprehension exercise the reader needs to be active, evaluate the text, can foresee the events in the text, be able to reread for better understanding and finding inconsistencies, assess his/her comprehension; he/she also need to be able to use his/her prior knowledge and monitor his comprehension.

The main goal of reading is reading comprehension, especially functional literacy tasks. Reading comprehension, which is a task that comprises many skills, describes outcome of taking out the meaning from a written text by using
one’s intellect. Curtis (2002) believes that there are a number of skills that the reader needs to employ in order to achieve maximum reading comprehension, skills such as: deciding about the main idea of the reading text, making questions regarding the content of the text and being able to answer those questions by employing context clues, and summarizing the passage.

Block and Pressley (2002) state that reading comprehension is usually considered as a process in which several elements are integrated. These elements are the ability to decode written materials, reader’s prior knowledge of the text, his/her vocabulary knowledge, and the reading strategies to comprehend the text. According to them, “comprehension involves more than 30 cognitive and metacognitive processes including clarifying meaning, summarizing, drawing inferences, predicting, and so on” (Block & Pressley, 2002). Moreover, Trabasso and Bouchard (2002, p. 177) assert that, “Comprehension strategies are specific, learned procedures that foster active, competent, self-regulated, and intentional reading”

There are a lot of studies that support the decisive role of vocabulary knowledge in a successful reading comprehension, especially in older children. For example, Ouellette & Beers (2010) have recently pointed out the important role oral vocabulary plays (contrasted with decoding skills) in foreseeing the success of Grade 6 students in reading comprehension.

Ouellette (2006), in a study involving Grade 4 students, found out that 28.5% of reading comprehension variance was the result of the breadth and depth of vocabulary knowledge. More specifically, the weight of vocabulary depth was heavier than those of vocabulary breadth and word recognition. On the contrary, Tannenbaum (2006) reported the importance of vocabulary breadth over vocabulary depth in reading comprehension. Anyhow, vocabulary breadth and depth had a significant relationship and in more than 50% of the cases the comprehension variance was common between the two.

As Nation (2001) stated, quite a few studies have shown that there exist a high correlation between the amount of vocabulary known reading comprehension. Stahl (2003), in one of her studies, maintains that readability studies formulae have revealed that the difficulty of the words in a text is the most important element in determining the difficulty of the text itself. Therefore, the size of one’s vocabulary knowledge is a strong predictor of one’s ability in reading comprehension.

**Lexical Inferencing**

According to Morrison (1996), lexical inferencing is using the available linguistic cues as well as other key points in a text in order to guess the meaning on unfamiliar word. Inferencing, in the opinion of Paribaht and Wesche (1999, p. 198), is a cognitive process that employs “familiar attributes and contexts” to distinguish unfamiliar elements in reading. Moreover, Haastrup (1991) states that lexical inferencing denoted a process of guessing the meaning of an unknown word by employing all linguistic cues available in the text together with the reader’s world knowledge, his/her linguistic knowledge, and his/her awareness of the context. Furthermore, “well-elaborated semantic knowledge, which includes developing knowledge of usage, collocations and other lexico-grammatical characteristics” is mainly achieved by learning new words through utilizing lexical inferencing in reading (Hunt, & Beglar, 2005, p. 28).

Advanced Students studying in English or other fields of study almost always are concerned with reading different written books and articles in English. Mostly these students suffer from lack of English vocabulary knowledge and this shortage is more problematic when it is impossible for them to access a dictionary and look that word up or to find someone with enough vocabulary knowledge and ask him/her. In such a case a thing which is very important and sometimes very useful is lexical inferencing. By lexical inferencing it is meant that a hypothesis about the meaning of a word is made actively and creatively and it is tested, if it is correct, it will enhance the accuracy and comprehension of the text otherwise it will lead to misunderstanding and misinterpretation of the text. It is interesting to note that sometimes vocabulary acquisition can occur through lexical inferencing because in this way a relationship can be imagined and made between the written word, the inferred meaning, and even the context of occurrence, and all these processes will lead to a complex interaction which can associate and join those word and meaning to one’s vocabulary knowledge. According to above mentioned ideas it can be said that to expose and learn less frequent words, one of the most useful things especially for advanced L2 learners is reading comprehension texts and one of the most recommended ways is applying lexical inferencing method while reading those texts (Huckin&Coady 1999).

Based on Haastrup’s definition in (1991), lexical inferencing is the application of linguistic and non linguistic hints included in a text to form “informed guesses” about the meaning of an unknown lexicon. She conducted a study about lexical inferencing in which she found that lots of various strategies namely, the internal structure of the word; top-down contextual and sentence-level clues were used by the participants. On the other hand, in her study, she introduced taxonomy for inference consisting of contextual, intralingual, and interlingual cues, and the results showed that those participants who used a combination of inferences and those who paid more attention to available information provided for them, were more successful.

So far several researches have shown that many factors contribute in successful lexical inferencing, some of them are listed below:

1) The essence and area of the word and text (Paribakht&Wesche, 1999).

2) The amount of the attention which is paid to the details in the text by the learner and his/her background knowledge about the possible meaning of the word (Frantzen, 2003).
3) The amount of learners’ basic pre-existing knowledge (Nassaji, 2006).

**Studies Done in the Domain of Lexical Inferencing**

During the last four decades several researchers have studied about the ways that L2 learners try to cope with unfamiliar lexical (Haastrup, 1991; Morrison, 1996; Nassaji, 2006; Babaei & Riazi, 2008).

Schmitt and McCarthy (1997) introduced three categories of knowledge sources which are used during lexical inference and played effective role in contextual understanding. These sources are linguistic, world, and strategic knowledge. Some of the recent studies also have confirmed the effectiveness of these knowledge sources. Nassaji (2006) is one of those researchers who have confirmed that the cooperation of several knowledge sources and strategies will lead to a more successful lexical inference. Moreover, he suggested an inference model which was able to distinguish strategies and the effective and appropriate use of them in association with different sources of knowledge.

Morisson (1996) is another researcher who studied lexical inferencing. In her study 20 L2 (French) learners read an authentic French text about the behaviors of pedestrian in a university context. In her study she used pairs think-aloud protocols which showed significant differences between high and low proficiency learners. In her research, high proficiency learners were more successful at lexical inferencing and they also used more knowledge sources and more combinations of various sources, and it was revealed that high proficiency learners tended to use contextual hints especially intralingual sources. On the other hand, low proficiency learners used contextual cues and knowledge sources but less than high proficiency learners, and as Morris proposed low proficiency learners should develop their lexical knowledge through word-analysis and their ability to integrate this lexical and contextual knowledge.

**Statement of the Problem**

So far in many studies it has been revealed that the amount of vocabulary one knows will affect his/her ability for decoding and understanding reading texts, in fact it seems that the more vocabulary knowledge causes the better text comprehension (Bengeleil & Paribakht, 2004; Schmitt & McCarthy, 1997). According to all previously mentioned ideas there is obviously a relationship between lexical knowledge and text comprehension but one point which is still controversial is the question that how the amount and distribution of lexical knowledge can affect comprehension especially reading comprehension (Babaei & Riazi, 2008). Clearly to get nearer to the answer to this question, this point should be taken into consideration in real EFL classes.

**Research Question**

Does text density in terms of the number of unknown lexical items have relationship with success of L2 learners’ lexical inferencing?

**Null Hypothesis**

Text density in terms of the number of unknown lexical items have no relationship with success of L2 learners’ lexical inferencing?

**III. Method**

**Participants**

The participants of this study were 30 L2 learners students selected from 90 EFL senior university students in Islamic Azad Universities of Boroujerd and Arak. All of the participants were native speakers of Persian, and they used English as a foreign language for general purposes. Their age ranged between 20 and 30, but gender was not considered as a determining factor in selecting the participants. The proficiency level of the participants was determined through the administration of a Quick Placement Test. Those whose scores in the test were between 48-55 were considered as the advanced-level participants of this study. (As it is specified in QPT ranking chart).

**Instruments**

In this study different instruments were used for data collections:

*Proiciency Test*

The proficiency level of the students was determined by administering a Quick Placement Test (version 1, 2002). The test consists of two parts; part one contains 40 items; 5 questions about situations, 15 questions on cloze passages, 20 questions of completion type. The second part contains 20 questions; 10 questions on cloze passages and 10 questions of completion type. All questions are multiple-choice items.

*Lexical Inferencing Texts*

In order to perform the designed research, under the direct supervision of the supervisor and advisor of the research, two passages that were supposed to be suitable for the guessing task in this study were selected from English websites.

**Procedures**

The research at hand was performed in six sessions, the researcher utilized the two selected passages described in instruments section for the guessing task in this study:

In the first session, the researcher piloted the texts with 15 senior EFL university students in Islamic Azad University, Arak branch. The participants were asked to read each text and underline the unknown words (Chegeni & Tabatabaei, 2013, P.125).

In the second session, the piloting was done with 15 senior EFL university students in Islamic Azad University, Boroujerd branch. In this task, the words which were underlined in the first task, and also were new based on the idea of some professors given to the participants. They were asked to write their synonyms in English or Persian. The
researcher choose the words which were found to be unknown words to be used in the main study. (chegeni & tabatabaei, 2013,P.125).

In the third session, the vocabulary and the comprehension tests with the related alternatives and distracters for each text were made separately. Then, it was piloted with 15 senior EFL university students in Boroujerd Azad University to be ensured of the correct alternatives and distracters, the appropriate timing, the administration procedures for each test, during the main study (chegeni & tabatabaei, 2013, P.125).

In the fourth session, a Quick Placement Test (QPT, version 1, 2002) which includes two parts was administered to 90 EFL participants in Boroujerd Azad University and Islamic Azad University, Arak branch. After administering the test a total number of 30 students who had the same language proficiency level were selected for this study. (chegeni & tabatabaei, 2013, P.125).

In the fifth session, the text entitled Dreams with high density of unknown words including 357 words totally and 20 unknown words was given to the participants and were asked to read it carefully and choose the correct choice for each question. All of the new words were content words involving four nouns: alacrity, invective, Breakthrough, infraction, five adjectives: slightest, futile, fractious, pugnacious, pellucid, and eleven verb: hurl, dither, deprecate, fawn, pontificate, kick, grab, deride, construe, heal, tumble. This text consisted of 10 multiple choice comprehension questions and 15 multiple choice questions for unknown words.

In the last session, the texts entitled The Mini Problem with low density of unknown words was administered. The text with low density of unknown words including 291 words totally and 10 new words was given to the participants to choose the correct answers. All of the new words were content words involving one noun: fervor, four adjectives: palpable, magnificent, insurmountable, blah, one adverb: sulkily, three verbs: relegate, seethe, yell and an expression: par for the course. This text consisted of 6 multiple choice comprehension questions and 10 multiple choice questions for unknown words.

IV. RESULTS AND DISCUSSIONS

Before starting the statistical procedure, all the scores obtained from the tests employed (except for the OPT) by the participants in the two experimental groups, that is, low density, high density, were converted to the scale of 100 to come up with homogeneous scores because the highest scores for different tests were different. (chegeni & tabatabaei, 2013,P.126).

Results of the OPT

Before starting the main experiment in order to make sure that the two groups were homogeneous with regard to their English proficiency, the results of their OPT were compared through administering a one-way ANOVA. Table 4.1 shows the descriptive statistics and the results of the ANOVA for this test. (chegeni & tabatabaei, 2013,P.125).

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>SEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>15</td>
<td>80.00</td>
<td>8.251</td>
<td>2.130</td>
</tr>
<tr>
<td>High</td>
<td>15</td>
<td>69.87</td>
<td>7.689</td>
<td>1.985</td>
</tr>
</tbody>
</table>

Regarding the information available in Table 4.2, the two means are not the same, but it cannot be told if their differences are statistically significant. To find this out, an independent-sample t-test was used. Table 4.3 shows the result.

<table>
<thead>
<tr>
<th>t</th>
<th>df</th>
<th>p</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.480</td>
<td>28</td>
<td>.002</td>
<td>10.13</td>
</tr>
</tbody>
</table>
According to Table 4.3, the amount of t-observed (3.480) is significant at the probability level of .002 which denotes a statistically significant amount, since if the significance level is smaller than the alpha level (.05), we should reject the null hypothesis and say that the difference is statistically significant, but if it is more than the alpha level we should retain the null hypothesis. As a result, the null hypothesis stating that, "text density in terms of the number of unknown lexical items have no effect on L2 learners’ successful lexical inferencing" can safely be rejected.

Discussions for the Research Question
In this study an attempt was made to answer the following research question appropriately:

Does text density in terms of the number of unknown lexical items have relationship with success of L2 learners’ lexical inferencing?

According to the data gained from this study and the statistics employing an independent-sample t-test, in preceding section it was seen that the the amount of t-observed (3.480), degree of freedom (28) and p value is .002. Moreover, mean score of high density unfamiliar words (69.87) is smaller than of low density unfamiliar words which is (80.00). As a result, the null hypothesis stating that, "text density in terms of the number of unknown lexical items have no relationship with success of L2 learners’ lexical inferencing" can safely be rejected.

This conclusion is in line with some previous research about the effect of text density regarding unknown vocabularies on learners’ successful lexical inferencing as Liu and Nation (1985) stated that when using high and low density text, words in low-density texts are easier to guess. (Yuen Shen & Shi Wu, 2009, P.190) The findings of this research also support Laufer’s (1997) suggestion that high density of unknown words in reading texts results in the inability of the learners’ successful inferencing.

The results of this study are parallel with what the other researchers in the field gained as Nassaji (2006) who argue that depth of vocabulary knowledge made a significant contribution to inferential success. According to Fukkink and Block (2001) these findings support the hypothesis that lexical inference depends heavily on the richness of the learners’ semantic and conceptual system (Yuen Shen & Shi Wu, 2009, P.191)

The result obtained from the first question is also supported by other researchers including Sternberg, (1987), Haynes, (1993), Fraser, (1999); Paribakht and Wesche, (1999), and Dubin & Olstain, (1993) who conclude that contextual factors including the number of occurrences of the unknown word, the importance of the unknown word to text comprehension, the density of unknown words in the text, text length, comprehension task, and the existence of clear contextual cues influence success in guessing an appropriate meaning for an unknown word. (Paribakht,2005,P. 721).

V. CONCLUSION AND IMPLICATIONS
To sum up, this study dealt with exploring the influence of number and density of unknown words on lexical inferencing. In order for the findings of this study to be pedagogically valid and applicable, they must be first subjected to replication and empirical validation. It is then and only then that the results and findings can be generalized to other populations. It should be mentioned that the findings of this study could enrich the literature in the area of second language acquisition development (Shoushinon, 2013, P.42) especially Iranian EFL learners’ lexical inferencing abilities.

The findings of this research revealed that text density regarding unknown vocabularies have effect on success of L2 students’ lexical inferencing. In other words the guessing success of the students in the texts with low density of unknown words interpreted that the lower the number of unfamiliar words, the higher the available clues for the participants to use for inferring the correct meaning of those words. When students come across a text with high density of new words, these new words are more difficult for the students to guess because they enjoy clues which are new words by themselves and consequently may add to the confusion of the students in referring the right meaning.

REFERENCES


Nastaran Chegeni received her M.A. in TEFL from Islamic Azad University, Shahreza Branch. She has presented some papers in national conferences in Iran. She has also published a couple of articles in different national and international journals. This paper is based on her Master’s thesis.

Omid Tabatabaei is an assistant professor at the English department of the Islamic Azad University, Najafabad Branch, Iran and currently the Head of the English Department in that university. His areas of interest are testing and assessment, research methodologies, psycholinguistics, language acquisition, and syllabus design. He has published and presented papers in international conferences and journals.