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

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

Omid Tabatabaei, Ph.D  
English Department, Najafabad Branch, Islamic Azad University, IRAN
Abstract

Lexical inferencing means guessing the meanings of unfamiliar words in context using available linguistic and other cues (Morrison, 1996). This study was an attempt to investigate the relationship between distribution of lexical items and L2 learners’ reading comprehension achievement. To this end, 30 EFL learners were selected and assigned to two groups of 15: two distribution groups of high and low. Then, the distribution (high and low) group’s performance on lexical inferencing was compared. The results of the study revealed that the participants inferred the meaning of unknown words more successfully in the text with low distribution of unknown words than in the text with high distribution of unknown words. The results of the study and their pedagogical implications are discussed. EFL teachers, learners, and practitioners may avail from the findings of this study.

**Keywords:** Lexical inferencing, distribution of lexical items, reading comprehension achievement
1. Introduction

From the four skills, reading is a process which contributes to success in learning a foreign language and good language learners are considered to be good readers as well (Bialystok, 1983). Based on Paribakht and Wesche (1999), a beneficial reader can guess the meanings of unfamiliar words in a text, and there is a strong relationship between vocabulary knowledge and reading comprehension. One way to assist learners with unknown words in a reading text is to train them to use contextual clues for inferring the meaning of these words rather than being merely dependent heavily on dictionaries. Thus, based on Nation (2001), guessing meaning from context is considered a sub-skill of reading.

Read (2000) refers to initial evaluation of the unknown word in terms of its contribution to the general understanding of the texts in order to illustrate how learners deal with unknown words. If an unknown word is not thought of having a major impact on comprehension, it is normally ignored and if it is judged to have a great contribution to determining the meaning, a variety of strategies are used to disambiguate it.

It is totally accepted that insufficient vocabulary size is the critical factor which influences reading and even language learning. Despite the fact that, the EFL/ESL reader cannot memorize vocabulary and reach the native speaker’s vocabulary size, it cannot be claimed that the EFL/ESL reader is not able to grasp the written text (Laufer, 1989). Although there is a general consensus in recent literature that the more vocabulary students know, the better they can decode and understand what they read (Bengeleil & Paribakht, 2004; Schmitt & McCarthy, 1997), there is little agreement as to exactly how the number and distribution of lexical items would affect reading comprehension (Babaei & Riazi, 2008). This process should take place in real EFL classroom situations and affecting need to be investigated.

Concerning the importance of lexical inferencing in second language research, this study tries to find out the relationship between distribution of lexical items and reading comprehension achievement.
2. Review of the Related Literature

2.1. Theoretical Background

Definition of Reading

Reading in a second language (L2) is complex, dynamic and multidimensional (Alderson, 2000). This is primarily, based on Alderson (2000), because it involves interactions among the reader’s interlanguage competence (e.g., incomplete, fragmented or not fully-developed linguistic, strategic, discourse and sociolinguistic competence), personal characteristics (e.g., learning and cognitive style, gender, motivation and volition, socioeconomic status, educational levels) and external contexts (e.g., topics, text characteristics, reasons to read, stakes of reading, time constraints).

Reading is a laborious activity that learners often can choose to do or not to do, it also requires motivation. To put it another way because reading often involves choice, motivation is crucial to reading engagement (Pintrich & Schunk, 2003). Motivation theorists attempt to understand the choices that individuals make among different activities available to them and their effort and persistence at the activities they choose (Pintrich & Schunk, 2003; Wigfield & Eccles, 2002a). Even the reader with the strongest cognitive skills may not spend much time reading if he or she is not motivated to read. In reading literature, much of the work relevant to readers’ motivation has been framed in terms of attitudes toward reading.

Rycik (2005) states that reading to learn is the most typical stage of reading development in middle school. Students reading complex text with the support of explicit reading instruction have better understanding of how to comprehend texts in a multitude of formats (Rycik & Irvin, 2005; Vacca, 2002). The literature presents a growing consistence of evidence positively confirming a systematic approach to instruction in reading for students of all ages (Pressley, 2002; Vacca, 2002). In addition, the literature supports a continuation of coherent and consistent reading practice and instruction each day, within all content areas (Brown, Pressley, Van Meter, & Schuder, 1996; Reeves, 2004). A continuation of reading comprehension instruction will propose support for middle school curriculum challenges.
Theories Related to Lexical Inferencing

Some theories about the nature of reading put a great deal of emphasis on strategies of guessing and predicting in the 1960s, 1970s, and well into the 1980s. Reading, from that point of view, was labeled as a “psycholinguistic guessing game” (Goodman, 1967). Based on this top-down reading model, readers are active samplers of text who combine text context with their knowledge about writing, language, and the world generally in order to read more efficiently, using prediction to take short-cuts in bottom-up processing of letters and words (Goodman, 1967; Smith, 2003).

Nevertheless, other research on the growth of L1 reading indicates that such top-down processing is not enough for effective reading. Rapid, precise recognition of letters and words, that is bottom-up, more input-constrained processing, must be dominated before fluent reading can take place as Perfetti, Goldman, and Hogaboam (1979) purport. In the same line, Huckin and Bloch (1993) also argue that emphasis on context has seriously underplayed the problems learners might have in simple bottom-up processing (i.e. graphemic identification of words). Such problems, as claimed by Huckin and Bloch, are rather common among ESL learners whose L1 is Chinese, a language with orthography and lexis radically different from those of English (Huckin & Bloch, 1993).

Another model of the reading process which blends top-down strategies and bottom-up, or data-driven, strategies, the supposed interactive model of reading, has, more recently, been developed (Eskey, 1997). Based on the interactive theory, effective readers use both top-down and bottom-up processing, although they are not always aware of the latter.

2.2. Previous studies

Bengeleil and Paribakht (2004) probed the use of inferencing strategies in reading comprehension by seventeen EFL Arabic-speaking medical students. In their introspective study, learners read an authentic English text on environmental refugees.

In another study, Paribakht and Wesche (1999) conducted a research with ten students, learners of English. During lexical inferencing, students used extra-linguistic and linguistic sources of information. Extralinguistic information let in the readers’ world knowledge and linguistic information referred to various levels of readers’ linguistic knowledge such as knowledge of grammar, morphology, punctuation, discourse, homonymy and cognates. Readers most often
applied grammatical knowledge at the sentence level (35%), followed by morphological analysis of unfamiliar lexical items (15%) and the knowledge of the world (9%). In his study, Nassaji (2003) partly confirmed these findings as far as the type of knowledge used is concerned, but the frequency of the knowledge type use was somewhat different, as the students mostly used world knowledge (46.2%), followed by morphological knowledge (26.9%) and grammar knowledge (11.5%).

Qian (2004) carried out a study with a questionnaire on the frequency of use of lexical inferencing strategies. The outcomes propose that the readers followed a top-bottom approach, i.e. they firstly used the information at the discourse/text level, then at the paragraph level, then at the sentence level and in the end at the word level. All the same, it is interesting to note that Qian also noticed and analyzed the readers’ actual behavior during reading. The discovered behavior was somewhat different from the readers’ self-perception of strategy use and they showed more of a bottom-up approach to the text.

In another study, Babaei and Riazi (2008) carried out a research to uncover Iranian EFL female students' lexical inferencing and its relationship to their L2 proficiency and reading skill. The results of their study showed that elementary students resorted to all clues including contextual, intralingual and interlingual to guess the meaning of unknown words, while intermediate students depended upon contextual clues only and advanced learners made use of contextual and intralingual sources and also advanced students were more successful in making correct lexical inferences. They further concluded that lexical inferencing ability did not show any significant relationship with students' reading comprehension performance.

Istifci (2009) probed lexical inferencing of Turkish EFL learners to show what inferencing strategies the learners at intermediate and low-intermediate levels apply when they seek to guess the meaning of unknown words and the similarities and differences between these two levels. It was found that students at the two levels attempted to use dissimilar types of inferencing strategies. Even so, the percentage of correct inferences was higher at intermediate level than low-intermediate level.

2.3. Research Question

The study addressed the following research question:
Does the distribution of unknown words in a text have any effect on the enhancement of L2 learners’ lexical inferencing?

2.4. Research Null Hypothesis

According to the question cited above, the study particularly aimed at testing the following null hypothesis:
The distribution of unknown words in a text has no effect on the enhancement of L2 learners’ lexical inferencing.

3. Methodology

3.1. Participants

The participants of this study were 30 L2 students, aged between 20 to 30 who were chosen from 90 EFL senior university students at Islamic Azad Universities of Arak Boroujerd. They were native speakers of Persian and used English as a foreign language for general purposes.

3.2. Instruments

Instruments employed in this study to carry out the intended research are as following:

Quick Placement Test

To identify the proficiency level of the participants, a Quick Placement Test (QPT) was used. The test consisted of two parts; part one contained 40 items, the second part contained 20 questions.

Lexical Inferencing Text

A text with high distribution of new words in the first paragraph (13 new words) and low distribution of new words in the next two paragraphs (9 new words) was chosen followed by 28 multiple choice questions (18 vocabulary, and 10 reading comprehension questions, 11 vocabulary and 5 reading comprehension questions in the first paragraph and 7 vocabulary 5 reading comprehension questions in the next two paragraphs ) to show the effect of distribution of new words in the texts on learners’ lexical inferencing.

3.3. Procedure

The present research was done in five sessions. the researcher utilized the one selected passage described in instruments section for the guessing task in this study:
In session 1, the researcher piloted the text with fifteen senior EFL university students in Islamic Azad University, Arak branch. The participants were asked to read text and underline the words which were unknown to them.

In session 2, the piloting was done with fifteen other senior students in Islamic Azad University, Boroujerd branch. They were asked to write their synonyms in English or Farsi. The researcher choose the words which were found to be unknown words to be used in the main study.

In session 3, to be ensured of the correct alternatives and distracters, the appropriate timing, and the administration procedures for each test, the vocabulary and the comprehension tests with the related alternatives and distracters for text were made separately.

In session 4, a Quick Placement Test including two parts was administered to 90 EFL participants, out of whom a total number of 30 students who had the same language proficiency level were selected for this study.

In session 5, the text entitled City Girl with 429 words (22 unknown words) which were content words including four nouns: strawberry, bush, hurricane, labor, seven verbs: avert, chop, thrill, indoctrinate, splay, squat, chaperone and two adjectives entire, pregnant which were in the first paragraph, and 9 new words in the next two paragraphs including three verbs: creeping up, materialize, prep, one noun: lawn, one adverb: literally, and four adjectives: cursory, insidious, woolen, potted, were given to the participants to choose the correct answers. Their text consisted of 28 multiple choice questions altogether.

4. 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 distribution, and high distribution, were converted to the scale of 100 to come up with homogeneous scores because the highest scores for different tests were different.

4.1. 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.
It can be seen in Table 4.1 that the result of the ANOVA is not statistically significant ($F_{observed}(3, 56) = .940, p = .427$); as a result, it can be said that the two groups were homogeneous at the beginning of the experiment with regard to their English proficiency.

### 4.2. Addressing the Null Hypothesis

For the null hypothesis, the performances of the two distribution groups had to be compared to see if the hypothesis could be retained or not. Table 4.4 indicates the descriptive statistics for this comparison.

<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>75.00</td>
<td>13.411</td>
<td>3.463</td>
</tr>
<tr>
<td>High</td>
<td>15</td>
<td>61.67</td>
<td>10.791</td>
<td>2.786</td>
</tr>
</tbody>
</table>

As it can be seen in Table 4.4, the two means are different from each other. In order to find out whether or not this difference is statistically significant, another t-test was employed. Table 4.5 presents the results of this t-test.

### Table 4.4

*Descriptive Statistics for Distribution*

<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>75.00</td>
<td>13.411</td>
<td>3.463</td>
</tr>
<tr>
<td>High</td>
<td>15</td>
<td>61.67</td>
<td>10.791</td>
<td>2.786</td>
</tr>
</tbody>
</table>

*The Results of t-test for Distribution*
By studying the amount of $t$-observed (3.000), according to Table 4.5, it can be understood that this amount of $t$ is significant at the probability level of .006, which is statistically significant. Therefore, the null hypothesis which states that "the distribution of unknown words in a text have no effect on the enhancement of L2 learners’ lexical inferencing" can also be rejected.

4.3. Discussions

In this study an attempt was made to answer the following research question appropriately:

Does the distribution of unknown words in a text have any effect on the enhancement of L2 learners’ lexical inferencing?

Based on the results of the $t$-test administered and studying the amount of $t$-observed (3.000), it can be understood that this amount is significant at the probability level of .006, which is statistically significant. Since the level of significance here was .006 and it is smaller than alpha level (.05). Therefore, the null hypothesis which states that "the distribution of unknown words in a text have no effect on the enhancement of L2 learners’ lexical inferencing" can also be rejected.

This result is parallel with what another researcher in the field gained such as Nation (2001) who argues that the presence of high distribution of unknown words in a text may seriously hinder comprehension.

It can be said that a high distribution of unknown words may result in the inability of the participants to use the available clues to infer the correct meaning because If the clues to the unknown word are in words which are themselves unknown to the reader, there are no clues for the reader to be used by him/her (Laufer, 1997).

The result obtained is also supported by other researchers including (Laufer, 1992a; Qian, 1999; Nation, 2001; Stahl, 2003) who conclude that the vocabulary size is a strong predictor of reading comprehension and there is correlations between vocabulary and comprehension, because it will affect the distribution of unknown word in a text. In addition, if the learners do not have

<table>
<thead>
<tr>
<th>t</th>
<th>df</th>
<th>p</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.000</td>
<td>28</td>
<td>.006</td>
<td>13.33</td>
</tr>
</tbody>
</table>
sufficient vocabulary size and could not recognize most of words in the context, the new word meaning are hard to be decoded correctly.

5. Conclusions

This study examined the relationship between distribution of lexical items and reading comprehension achievement. As it was mentioned before, the findings of the study revealed that the distribution of unknown words in a text definitely affects learners’ successful lexical inferencing. Based on the results obtained through the statistical analysis on the collected data mentioned previously, it can be safely claimed that there is a significant difference between distribution of unknown words and lexical inferencing of Iranian EFL learners. The participants inferred the meaning of new words in the text with low distribution of unknown words more successfully than in the text with high distribution of unknown words. So, L2 learners will probably be more successful in guessing the meanings of the unknown words by including explicit strategy on contextual guessing in the reading and vocabulary instruction.

6. Suggestions for Further Research

The preliminary purpose of this study was to probe the existence or lack of existence of any interrelationship between of number and distribution of unknown words and lexical inferencing ability of Iranian language learners. However, in order to complement the findings of the present study, the topic needs to be further explored in some other studies.

So further research may be necessary in the following areas:

One area could be examining the relationship of the same independent variables (lexical inferencing) and other language skills such as speaking, listening, or even writing performances.

Furthermore making a comparison between the results of the present study and also the same ones carried out in EFL contexts with those done in ESL contexts can be another area of study.

Finally different age groups and proficiency levels with diverse educational backgrounds can be the subjects of this study to see if they come up with the same results.
References


Nastaran Chegeni & Omid Tabatabaei: Lexical Inferencing: The Relationship between Distribution of Lexical Items and L2 Learners’ Reading Comprehension Achievement Learners

(Eds.), Second language vocabulary acquisition: A rationale for pedagogy (pp. 20-34). New York: Cambridge University Press.


In P. Bogaards & B. Laufer (Eds.), *Vocabulary is a second language: Selection, acquisition, and testing* (pp.155-169). Amsterdam: John Benjamins Publish Company.


