Title

The Impact of Text Length on EFL Students’ Reading Comprehension

Author

Saeed Mehrpour
Abdolmehdi Riazi, Ph.D

Department of Foreign Languages and Linguistics
Shiraz University

Bio Data

Brief history of professional expertise: Abdolmehdi Riazi is associate professor of TESL and teachers graduate courses (master’s and Ph.D.) in Applied Linguistics at Shiraz University.

Saeed Mehrpour is a Ph.D. candidate in TEFL in the department of Foreign Languages & Linguistics of Shiraz University. He is currently writing his dissertation.

Qualifications:

Abdolmehdi Riazi holds a Ph.D. in TESL from University of Toronto, Canada.
Saeed Mehrpour holds an MA in TEFL and is a Ph.D. candidate.
ABSTRACT

This study investigated the impact of text length on reading comprehension. Three reading comprehension passages accompanied by 30 items were reduced to two thirds of their original lengths. The truncated versions along with the original version were randomly administered to 100 male and female college students of English (50) and non-English majors (50). Although the participants performed marginally better on the shortened version of the test, a two-way ANOVA revealed that the length of the text did not have any statistically significant effect on the students' performance on the two versions of the test. No interaction between the length of the test and the students' major was found either. Further analysis of the data indicated that sex did not have any significant effect on the performance of the students on the two versions of the test. The article concludes with suggestions for further investigations into the issue in question.

Introduction

Broadly speaking, factors affecting reading comprehension can be classified into two general categories: (1) reader variables, and (2) text variables. The first category of variables which include factors such as reader's background knowledge, reader's skills and abilities, and reader's motivation and attitude are internal to readers. The second category of variables which include factors such as text content, text type and genre, text organization, text typographical features, and text readability are internal to texts rather than to readers (Alderson, 2000).

The notion of comprehensibility of a text is closely related to a most familiar notion in reading comprehension, namely, readability. According to Richards et al. (1992, p. 306), “readability refers to how easily written materials can be read and understood. It depends on many factors including (a) the average length of sentences in a passage, (b) the number of new words a passage contains, and (3) the grammatical complexity of the language used.” The notion of readability is, of course, a controversial issue since researchers have found that linguistic complexity may not be detrimental to comprehension. Bernhardt (1984 cited in Chastain, 1988, p. 232), based on research findings, stated that “syntactic simplicity may decrease text cohesion and thereby hinder comprehension.”

Shifting views towards reading comprehension as an interactive process between the reader and the text which involves the reader’s activating relevant schemata to recreate the writer’s meaning has led to a different reading model which, in turn, uses different criteria to measure readability. Based on this model, the reader’s background knowledge and interest are the two most important criteria which determine the readability of a text to a great extent (Chastain, 1988). It is now generally believed that a range of reader factors such as motivation, background knowledge, and previous reading experience affect the reading process to a considerable extent (Fulcher, 1997).
Another factor which can contribute to the level of difficulty of a passage is its length. In fact, it is an intuitive belief that the longer a passage, the more difficult it will be. With regard to this point, Chastain (1988, p. 234) contends that “language teachers usually favor short reading passages. This tendency may have its roots in a desire not to overload the students, but it probably is an outgrowth of the idea of reading as a laborious process of deciphering a complex and sometimes unfamiliar linguistic code for which the students were required to know all grammar and vocabulary. The result has been that students focus too much attention on language as opposed to meaning, and they read much too slowly.”

Based on psychological approaches to reading comprehension, reading theorists recommend that students read longer passages. Stanovich’s (1980) interactive-compensatory model is an example of such psychological approaches. This model indicates that the reader processes texts by using information provided simultaneously from several sources which can compensate for deficiencies at one level by drawing on knowledge at the other (either higher or lower) levels. These sources of information comprise all those considered separately in skills model of reading comprehension which draws on inside-the-head or top-down processes and psychological model of reading comprehension which makes use of inside-the-head or top-down processes. Stanovich’s model implies that when a reader can not comprehend a text based on inside-the-head or top-down processing, (s)he can resort to other sources of information such as text-based information. A second closely related implication is that the more text-based information the reader has access to, the easier the comprehension of the content of the text will be. Anderson (1984, p. 94) supports this position, saying that “research on reading miscues has demonstrated that short items are harder to read than longer ones because reading involves building up expectations on the basis of redundancies.” Likewise, Feldmann and Stammer (1987, p.255) assert that, “the more clues the learners are able to pick up, because of the natural redundancy of a text, the more developed is their foreign language competence and the better they will accomplish the task.”

With respect to the relationship between text length and reading comprehension ability, Alderson (2000, p. 108) maintains that, “a problem all reading-test developers face is how long the texts should be on which they base their tests. Text length is a surprisingly under researched area.”

In testing any language ability, one should bear in mind that an individual’s performance on a test is highly determined by both his/her language ability as well as test method facets. Broadly speaking, test method facets include aspects of a test such as rubrics, testing environment, and input (Bachman, 1990). According to Bachman & Palmer (1996, p. 52), “the input format of a test encompasses aspects such as channel, form, language, length, type, degree of speediness, and vehicle. The length of the input, for example, influences the amount of interpretation required. Input can be presented in very short chunks and thus require limited interpretation, or it can be presented in extended discourse requiring more extensive interpretation”. Alderson, (2000, p. 153-154) also maintains that, “the length of input is clearly very important in reading. Texts may be a single word, a phrase (as in the case of public warnings and notices), single sentences (in advertisements), or they may be whole novels, 20-page academic articles, two-page memoranda, or two paragraphs on a
postcard.” He further states that, “whilst length is related to the amount of interpretation needed, this is not necessarily direct: a post card may be very enigmatic, a novel very explicit” (p. 154).

**Statement of the problem**

In connection with the effect of the length of a passage on reading comprehension, one can pose the following research question: ‘does the length of a passage affect its level of comprehensibility? Another relevant research question is whether language learners learning English as their major perform differently on such passages as compared to non-major English language learners.

The purpose of this study was, thus, to investigate one of the characteristics of the test input, i.e., length on readers’ performance. The findings of the study might shed more light on the relationship between the length of passages in tests of reading comprehension and their comprehensibility. The results of the study might be of use in selecting reading materials for teaching and testing purposes. They may also provide information which might be of value for reading comprehension teaching purposes.

**Review of the related literature**

With regard to the effect of different types of modification on reading comprehension texts, a number of studies are cited in the literature the majority of which have come up with mixed results. One of the first studies conducted in this respect is the one done by Newsom and Gaite (1971). This study was designed to investigate the learning and retention of prose materials. Immediate and one-week retention was examined in subjects who had either read a long 2300-word passage (LP) on science fiction or a 300-word short passage (SP). The short passage was derived from the information retained over a one-week period by subjects who had read and learned the long passage in a pilot study. The results obtained from the analysis of data revealed that subjects who read the short passage scored significantly higher on the one-week retention test than subjects who read the long passage. There was no difference in the mean scores of the two groups on immediate retention. It was also found that retention over one week was greater for subjects who learned the short specially prepared and organized materials than for the subjects who learned longer prose materials which contained much detail and narrative descriptions.

Engineer (1977 cited in Alderson, 2000) found that when texts longer than 1,000 words were used, the abilities that could be measured changed. This, in turn, suggests that longer texts allow testers to assess more study-based abilities and to reduce reliance on sentential processing abilities that might measure syntactic and lexical knowledge more than discourse abilities.

Terzi (1984) investigated the impact of text length on the development of second-language reading skills. On the basis of positive results obtained using a novel by Isaac Asimov in the early stages of an instrumental English reading course for Brazilian electrical engineering students, she questioned the soundness of the common practice of utilizing shorter texts at the initial stages of second-language instruction and reserving longer texts for advanced students. She finally argued that different texts invite distinct reading
strategies partially as a function of their length. Longer texts appear to require a more intense interaction, which encourages increased sensitivity to the linguistic aspects of the written language.

Strouhal et al. (1986) discussed psycholinguistic, sociolinguistic, and text-pragmatic aspects of text comprehension along with the empirical findings from educated and non-educated students (N = 130) who were interviewed and administered cloze tests to determine their ability to comprehend original and modified versions of a legal text. Based on their study, they found that although it is possible to modify specialized texts to make them more accessible to the general public, there are certain limits beyond which such modifications can not go before the text’s legal status is lost.

Strother and Ulijn (1987 cited in Alderson, 2000) compared reading comprehension scores of native and non-native subjects reading original texts and texts which had been simplified syntactically but not lexically. They found no differences in the performance of the two groups of subjects and concluded that simplifying syntax does not necessarily lead to more readable texts, since a thorough syntactic analysis of a text may be unnecessary while reading the text for reading comprehension purposes.

In another study, Campbell (1987) examined two adaptations to explore the claim that ‘simplification’ always create a text which is easier to read, and also to see if any loss in narrative interest is incurred through simplification. The type of simplification discussed by her was that of lexical simplification since lexis is believed to be the aspect of reading texts which is most regularly identified by readers as difficult. By providing two examples she claimed that lexical control showed that simplification may in some cases render a text more difficult to process, because criteria which make native speakers of English judge certain words as more difficult than others may not be relevant to certain groups of learners. Based on her comparison of the original and adapted forms of a story, she demonstrated that character contrasts, where these are linguistically determined, are inevitably weakened by lexical simplification, with the result that the narrative becomes less interesting.

Abrahamsen and Shelton (1989) carried out a study in which ninety-two adolescents with learning disabilities were randomly assigned to four groups to determine the effects of semantic and syntactic complexity on the reading comprehension of content area prose. One group served as a control group and read a social studies passage without simplification. The three treatment groups read passages with syntactic and/or semantic modifications. The results of the study indicated that comprehension was significantly better for those groups who read passages with combined semantic and syntactic modifications and syntactic modifications alone, when they were compared to the control group. It was also found that semantic modifications alone did not improve comprehension significantly.

Susser and Robb (1990) pointed out some important problems with respect to the use of graded readers in Extensive Reading Courses. As an example of such problems they claim that the process of simplification which is widely used in preparing graded readers leaves
writing that is more difficult to understand than the original, because, for instance, cohesion, coherence, and discourse structure are impaired.

In still another study, Yano et al. (1994) tested the hypothesis that elaborative modification in oral foreigner talk, where redundancy and explicitness compensate for unknown linguistic items, offers an alternative approach to written text modification. In their study thirteen reading passages were presented to Japanese college students (N = 438) in one of three forms: native baseline, simplified, or elaborated. The results obtained from the analysis of the data showed that comprehension was highest for the simplified version, but was not significantly different from the elaborated version. It was also revealed that modification type interacted significantly with the kind of test item used to assess three different types of comprehension processes--replication, synthesis, and inference--suggesting that different kinds of text modification facilitated different levels of comprehension.

Oh (2000) carried out a study relatively similar to that of Yano et al. (1994). In this study the researcher investigated the relative effects of two types of input modification--simplification and elaboration--on Korean high school students’ EFL reading comprehension. Six English reading passages in one of three forms--(a) baseline, (b) simplified, or (c) elaborated--were presented to 180 students, who were divided into two proficiency levels (i.e., high proficiency and low proficiency). Comprehension was assessed by an 18-item multiple-choice test, which included items for assessing (a) general, (b) specific, and (c) inferential comprehension. In addition, students’ perceptions of their comprehension were measured by their responses on a 6-point unipolar scale. The test data were analyzed by a 2-by-3 analysis of variance. The results of the study support the suggestion that input should be modified in the direction of elaboration rather than by artificial simplification, because elaboration retains more native-like qualities than and is at least equally successful as simplification in improving comprehension. The researcher suggests that instruction with elaborated input should accelerate the progression to fluent reading of unmodified materials, which is the ultimate goal of foreign language reading instruction.

Considering the above reported studies, one can come to the point that the results of the studies are not consistent with each other and are to some extent mixed. This fact indicates the necessity for further research in this area to fill the gap on the relationship between the effect of different types of text modification, of which text length reduction is an example, on comprehensibility of texts by language learners.

**Method**

**Participants**

In order to conduct the study, 100 advanced male (48) and female (52) language learners of whom 50 were majoring in English as a foreign language (English-major students) and the other 50 majoring in other disciplines at one of the major universities in Iran took part in the study. The students’ age ranged from 21 to 30.

**Instruments**
Two reading comprehension tests were utilized in the study. One of these two tests was, in fact, a shortened version of the original test in which the text length was reduced to two third of the original text. The original version of the test was taken from the TOEFL practice tests (ETS, 1995). Each test was composed of three passages accompanied by 30 multiple-choice items. The passages were on different topics including sociology, medicine, and geology. With respect to the procedures followed in shortening the passages, it should be stated that there are a number of criteria mentioned in the literature for modification of texts such as the ones stated by Aebersold and Field (1998), and Yano et al. (1994). A passage can, in fact, be modified in two different ways. It can either be simplified using more frequent, general usage vocabulary and less complex grammar structures or elaborated using more complex grammar structures, more difficult vocabulary, and adding background or cultural information.

The type of modifications made to the passages in this study were, to some extent, different from those already mentioned. In fact, the criteria used in shortening the passages were new to some degree and were more in line with those employed by Courchene (1995) in the technique known as summary cloze. Broadly speaking, in this technique the content of the original passage is summarized into a text about one third the length. The shortened text is then changed into a rational cloze passage. In summarizing a passage through this technique, all the key information is preserved and redundant or insignificant information is deleted. The passage is then rewritten using different wordings.

Considering the points already noted, the researchers used two main criteria in shortening the passages. The first criterion was the significance of the pieces of information presented in the body of each passage appearing in the two tests and their degree of contribution to the overall understanding of the passages. Following this criterion, we removed from the passage any pieces of information which were thought of as redundant and insignificant. The second criterion used for text length reduction was removing grammatical redundancies from the passage by changing clauses into phrases where possible and deleting elements which were grammatically recoverable such as changing long relative clauses into reduced clauses. An attempt was made to maintain, to a great extent, the totality of the information presented in the original passages while shortening them. By applying these two text length reduction criteria the passages were reduced from 346, 333, and 371-word passages to 242, 236, and 264-word long passages respectively.

The reliability of the two versions of the test were calculated using KR-21 formula of estimating reliability. The indexes obtained were .48 for the shortened version of the test and .62 for the original version. As for the validity of the tests, since the original version of the test was taken from ETS TOEFL practice test, the validity of the test was taken for granted.

**Procedure**
The two versions of the test were administered to the two groups (English majors and non-English majors) of students participating in the study. Half of the students in each group received the original version of the test and the other half received the shortened version randomly.
Analysis of data
After the administration of the tests, the data were collected and subjected to statistical analyses. Considering the fact that the two versions of the test were administered to the two groups of subjects, that is, those learning English as their majors and non-major learners of English four different types of raw scores could be identified and classified: scores belonging to majors taking the shortened version of the test (Major/Short), scores belonging to majors taking the original or the long version of the test (Major/Long), scores belonging to non-majors taking the shortened version of the test (Non-major/Short), and finally those belonging to the non-majors taking the original or the long version of the test (Non-major/Long). The descriptive statistics were calculated for the two versions of the test and the four sets of raw scores. To answer the two main research questions, the data collected were subjected to a two-way analysis of variance. As an after-thought and in order to see whether the variable sex had any impact on the performance of the two groups of subjects on the two versions of the test, independent t-test was used as well.

Results and discussion
Table 1 presents the descriptive statistics for the two versions of the tests, namely, the shortened and the original regardless of the students’ field of study, and also the reliability indexes estimated for them.

<table>
<thead>
<tr>
<th>Test</th>
<th>Mean</th>
<th>SD</th>
<th>Max</th>
<th>Min</th>
<th>KR-21</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short</td>
<td>20.70</td>
<td>3.46</td>
<td>29</td>
<td>13</td>
<td>.48</td>
</tr>
<tr>
<td>Original</td>
<td>20.22</td>
<td>4.04</td>
<td>29</td>
<td>12</td>
<td>.62</td>
</tr>
</tbody>
</table>

As can be seen from the above table, the means obtained for the short and the original version of the test (20.70 and 20.22, respectively) are quite close to each other. In order to see whether the marginal difference existing between the means obtained for the short and original version of the test was statistically significant, we ran an independent t-test. The following table shows the result obtained from this statistical analysis.

<table>
<thead>
<tr>
<th>Variable</th>
<th>No of cases</th>
<th>Mean</th>
<th>SD</th>
<th>SE of Mean</th>
<th>Df</th>
<th>2-tail Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short</td>
<td>52</td>
<td>20.70</td>
<td>3.46</td>
<td>.976</td>
<td>98</td>
<td>.617</td>
</tr>
<tr>
<td>Original</td>
<td>48</td>
<td>20.22</td>
<td>4.04</td>
<td>1.060</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The result of the above analysis reveals that the difference between the means of the two sets of scores on the two versions of the test is not statistically significant.

The descriptive statistics for the four sets of raw scores were also computed. The following table summarizes these indexes.
Table 3: Descriptive statistics for the four sets of raw scores

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean</th>
<th>SD</th>
<th>Max</th>
<th>Min</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS</td>
<td>20.80</td>
<td>3.797</td>
<td>29</td>
<td>13</td>
</tr>
<tr>
<td>ML</td>
<td>20.08</td>
<td>4.443</td>
<td>29</td>
<td>13</td>
</tr>
<tr>
<td>NS</td>
<td>20.60</td>
<td>3.162</td>
<td>26</td>
<td>13</td>
</tr>
<tr>
<td>NL</td>
<td>20.36</td>
<td>3.684</td>
<td>26</td>
<td>12</td>
</tr>
</tbody>
</table>

MS = Major/Short, ML = Major/Long, NS = Non-major/Short, NL = Non-major/Long

The table shows that the means of the four sets of scores are quite close to one another. To see if there existed any interaction between English major students and non-English majors and the type of test they took, we ran a two-way analysis of variance (2-way ANOVA). The results obtained from this analysis are displayed in Table 4 below.

Table 4: Two-way analysis of variance for the effect of test length and the students’ field of study and their interactions

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>DF</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig of F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main Effects</td>
<td>5.800</td>
<td>2</td>
<td>2.900</td>
<td>.201</td>
<td>.818</td>
</tr>
<tr>
<td>Group</td>
<td>.040</td>
<td>1</td>
<td>.040</td>
<td>.003</td>
<td>.958</td>
</tr>
<tr>
<td>Test Length</td>
<td>5.760</td>
<td>1</td>
<td>5.760</td>
<td>.399</td>
<td>.529</td>
</tr>
<tr>
<td>2-way Interaction</td>
<td>1.440</td>
<td>1</td>
<td>1.440</td>
<td>.100</td>
<td>.753</td>
</tr>
<tr>
<td>Group / Test Length</td>
<td>1.440</td>
<td>1</td>
<td>1.440</td>
<td>.100</td>
<td>.753</td>
</tr>
<tr>
<td>Explained</td>
<td>7.240</td>
<td>3</td>
<td>2.413</td>
<td>.167</td>
<td>.918</td>
</tr>
<tr>
<td>Residual</td>
<td>1385.600</td>
<td>96</td>
<td>14.433</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As the table indicates, there was no statistically significant difference between the performances of the students on the two versions of the test. No interaction between the English major and non-English majors (indicated as 'group' in the above table) and the type of test was observed. A possible reason for the non-significant effect of the test length on the performance of the students participating in this study is that the students, as stated above, were all advanced language learners who had been learning the language formally for at least 4 years. Such being the case, the lengths of the passages included in the test which were reduced to almost two thirds of their original lengths had not affected the students’ reading performance significantly. Alderson (2000, p.65) lends further support to this finding saying that, “text variables only have a crucial role when materials are conceptually more difficult or unfamiliar and when readers are relatively less able” (emphasis added). In fact, the results might have been different if the two groups of students were of different levels of proficiency.

In order to gain further support for the above-noted finding and to see whether shortening the texts had actually made them simpler, we calculated the readability indexes for both the original and the shortened passages appearing in the two versions of the tests using the Flesch Reading Ease formula. This formula, which is one of the readability formulas
Currently used, computes readability based on the average number of syllables per word and the average number of words per sentence. Scores range from 0 to 100. The higher the score, the greater the number of people who can readily understand the text (Fulcher, 1997). The following table depicts the readability indexes or difficulty estimates obtained for the passages in the two tests.

Table 5: Readability indexes obtained for the passages

<table>
<thead>
<tr>
<th>Test</th>
<th>Passage # 1</th>
<th>Passage # 2</th>
<th>Passage # 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original</td>
<td>63.4</td>
<td>54.8</td>
<td>35.2</td>
</tr>
<tr>
<td>Shortened</td>
<td>59.8</td>
<td>52.1</td>
<td>39</td>
</tr>
</tbody>
</table>

Based on this table, we come to the point that the shortened versions of the passages turned out to be more difficult than the original passages in two cases (passages # 1 and 2), though this difference was not very high, and only in the case of the third passage it led to an increase in the readability of the passage. This being so, the students must have had a better performance on the original version of the test than its shortened version. However, as the comparison of the means obtained for the performance of the students on the two versions of the test and the four sets of raw scores revealed, in all cases the students had a better performance on the shortened version of the test as compared to the original version, though the difference was very marginal. These contradictory findings render support to the issue of crudity of readability indexes and the fact that readability can not be taken as a good measure for determining the difficulty of texts (Chastain, 1988). As noted already, the Flesch formula is a function of the length of sentences and the size of words in a text, and nothing else. As Fulcher (1997, p. 501) states, “these factors are not, in themselves, what makes a text difficult or easy; but they are predictors of text difficulty.” In fact, important text variables, such as specialist use of lexis, conceptual and propositional complexity, textual organization, and syntax are not taken into account in estimating readability through Flesch formula (Fulcher, 1997). And, as mentioned previously, there are more important factors such as reader factors which contribute to the comprehensibility of texts.

Broadly speaking, the findings of this study are in direct line with Yano et al. (1994) study which indicated that comprehension was highest for the simplified version of the test, but was not significantly different from the elaborate version. These findings are also in support of Strother & Ulijn (1987), Campbell (1987), Susser & Robb (1990), and Oh (2000) who contend that modification does not necessarily result in better comprehension on the part of language learners. They also support Alderson (2000), as quoted above.

So far the answers have been provided for the two research questions posed in this study. With respect to the students’ sex, (52 females vs. 48 males), we decided to see if any relationship existed between sex and performance on the two versions of the test. Result of the analysis showed that the difference between the mean scores of male and female students were not statistically significant (mean= 21.08, 20.15; SD= 3.423, 2.968, respectively; 2-tail significance = .475). The result of another independent t-test revealed that the difference between the mean scores of these students on the original version of the test was not statistically significant, either (mean= 20.91, 19.84; SD= 3.088, 4.220, respectively; two-tail significance = .480).
Similar analyses were done to the data obtained from the performance of male and female English majors taking the two versions of the test. The result of an independent t-test indicated that the difference between the mean scores of male and female English majors on the shortened version of the test was not statistically significant (mean= 21.41, 19.80; SD= 3.965, 4.109, respectively; two-tail significance = .312), nor was any statistically significant difference was found (mean= 19.69, 20.54; SD= 3.313, 3.616, respectively; two-tail significance = .657) for their performance on the original test.

Conclusion
This study was carried out to investigate one aspect of input, i.e., length in reading comprehension tests. This aspect was explored through manipulating text length and its effect on EFL students’ reading comprehension performance. With regard to the first research question, that is, whether the performance of the language learners differed when they performed on passages with roughly similar contents but in different lengths, the results of the study suggest that reducing text length affects performance on reading comprehension only very marginally, at least when advanced language learners are involved. In connection to the second research question, namely, whether language learners learning English as their major perform differently on original passages and their reduced form as compared to non-major English language learners, no interaction was observed between the English and non-English major students and the type of test they took (test with original passages vs. test with shortened passages). As for the last research question, i.e., whether the variable sex had any effect on the performance of the students on the two versions of the test, the results indicated that this factor did not have any significant effect on the performance of the students. Based on these findings, we can maintain that language teachers do not have to limit themselves to using short passages or simplified readers when they teach and test reading comprehension.

References


Terzi, S. B. (1984). A extensao do texto como determinante de mudancas de postura frente a leitura: uma experiencia (Text length as a determiner of attitude changes regarding