

Content schemata, linguistic simplification, and EFL readers' comprehension and recall

Mohammad Hossein Keshavarz and Mahmoud Reza Atai
Tarbiat Moallem University
Iran

Hossein Ahmadi
Azad University
Iran

Abstract

This study investigated the effects of linguistic simplification and content schemata on reading comprehension and recall. The participants, 240 Iranian male students of English as a foreign language (EFL), were divided into 4 homogeneous groups, each consisting of 60 participants (30 with high proficiency and 30 with low proficiency). To elicit data, the study used 2 types of texts: content-familiar and content-unfamiliar. Each type appeared in 4 versions: original, syntactically simplified, lexically simplified, and syntactically-lexically simplified. Each participant group was tested on 1 of the linguistic versions of the content-familiar and content-unfamiliar texts. Data analyses showed a significant effect of the content and EFL proficiency, but not of the linguistic simplification, on reading comprehension and recall. The effect of the linguistic simplification on reading comprehension and recall is interpreted in the light of the interaction of content and linguistic simplification.

Keywords: EFL, reading comprehension, recall, content schemata, linguistic simplification, language proficiency

Effective teaching of reading comprehension necessitates an understanding and analysis of its nature and components, including both text and reader variables. Grabe (1997) put it in this way:

The central components of reading processing include: orthographic processing, phonological coding, word recognition (lexical access), working memory activation, sentence parsing, propositional integration, propositional text-model formation, text-model development, and the development of an appropriate situation model (mental model). (p. 9)

Although research on text factors, reader variables, and second language (L2) reading comprehension abounds in the literature, the overall picture of the patterns of interaction among

the major factors is far from clear (see Oh, 2001). More specifically, we need to understand when and why some readers fail to generate appropriate situation models of the text they read in spite of adaptation applied to the linguistic features (Grabe, 1997).

The present study investigated, from among the components of reading cited above, the contribution of vocabulary and syntactic knowledge, content and background knowledge (content schemata), and L2 proficiency (as a reader variable) to reading comprehension and recall. Considering the important role of the interactions among the aforementioned variables in making pedagogical decisions for L2 reading courses, the findings offer instant implications for materials developers and teachers of English as a foreign language (EFL).

Content Schemata and Linguistic Simplification

Prior knowledge plays a supportive role in comprehending a written message. The earliest study on the impact of schemata on reading comprehension dates back to the classical research of Bartlett (1932). In his study, English participants were asked to read and recall a story from an unfamiliar culture, and the major finding was that the recall was inaccurate. Distortions found in the retellings of the story conformed to the past experiences of the readers, and additions to and elaborations on the storyline in the retellings caused redundancies.

Johnson (1982) investigated the effect of the cultural origin of prose on the reading comprehension of Iranian intermediate and advanced students of English as a second language (ESL) at the university level. The results revealed that the cultural origin of the stories had a greater effect on comprehension than the syntactic or semantic complexity of the text. Some other studies have shown similar effects in that participants better comprehended or remembered passages that were more familiar to them (e.g., Ammon, 1987; Carrell, 1983; Johnson, 1982; Langer, Bartolome, Vasquez, & Lucas, 1990). The literature supports the position that content schemata have a greater role than language on reading comprehension.

With regard to linguistic simplification, Blau (1982) studied how manipulating the degree to which sentences were combined affected the reading comprehension of two groups of ESL students. The results indicated that lower readability levels did not facilitate reading comprehension. Blau argued that simplifying authentic discourse text may distort its naturalness. Questioning the validity and usefulness of linguistic simplification, Yano, Kong, and Ross (1994) found further evidence against a positive effect of simplification on reading comprehension. A study conducted by Parker and Chaudron (1987) also supports the idea that linguistic simplification does not necessarily make a text easier to understand as a whole. Byrd (2000) referred to the danger of using inauthentic simplified materials, stating that "these materials can remain difficult because of the loss of connectors and other language used to guide the reader through the text" (p. 2). Along the same line of research, Britton, Gulgoz, and Glynn (1993) demonstrated that presenting content in appropriate ways improves readability much more than the older traditions of text simplification. A crucial issue to investigate is the effect of linguistic (syntactic or lexical) simplification on reading performance across content-familiar and content-unfamiliar texts.

Content Schemata and EFL Proficiency

In addition to the role of content in EFL reading comprehension, the potential interaction between content and learners' proficiency also concerns researchers in EFL reading. Carrell (1984) suggested that the nonnative readers in her study failed to use background information because they were linguistically bound. The language itself required so much attention that the nonnative readers tended to process at the word and sentence levels and did not attend to the top-level organizational features and background information. She suggested that a threshold level of language proficiency may allow readers to engage in top-down processing. Similarly, Alderson (1984) found a threshold level of EFL proficiency above which the effect of EFL proficiency on reading comprehension and recall changed.

Carrell's (1983) study showed that higher proficiency students recalled more from content-unfamiliar texts than from content-familiar ones. Lee (1986) similarly found that his participants recalled more from content-unfamiliar texts than content-familiar ones. Studies conducted by Koh (1985) and Peretz and Shoham (1990) support the idea that participants do not necessarily perform their best on texts with familiar content. Clapham (1996) found a threshold score on IELTS above which her participants' performance on texts outside their discipline showed a great improvement. The question is whether this content and proficiency interaction remains the same in the case of reading comprehension, as measured by multiple-choice (MC) questions and recall.

Linguistic Simplification and EFL Proficiency

According to Widdowson (1979), linguistic simplification is supposed to "bring the language of the original within the scope of the learners' transitional linguistic competence" (p. 185). However, the literature on the interaction between modification (i.e., simplification in particular) and L2 proficiency indicates mixed results. For instance, Blau (1982) showed that simplified texts with lower readability levels, as measured by common readability formulae, did not facilitate the reading comprehension of his ESL participants. Strother and Ulijn (1987) compared native and nonnative speakers' reading comprehension of original texts versus their comprehension of texts simplified syntactically but not lexically. Their conclusion confirmed the earlier research finding: syntactic simplification does not necessarily make the text more readable. A possible interpretation might lie in the point that, as Alderson (2000) maintains, syntactic adaptation of a text may distort the message and lower its authenticity value (Davies, 1984; Widdowson, 1978). For example, Mountford (1975) concluded that the simplification of a scientific research article resulted in a change in the illocutionary force of the text. Davies (1984) came up with a text simpler than an original as measured by readability formulae and cloze tests. However, as is echoed in the more current reading literature (e.g., Alderson, 2000; Fulcher, 1997), the validity of the readability formulae have been questioned because they do not measure more than the superficial aspects of text difficulty. Therefore, readability predictions should provide an unbiased estimate of text difficulty by considering a wider range of text variables including topic, cohesion, coherence, and reader factors (e.g., readers' topical and background knowledge).

Oh (2001) investigated the relative effects of simplification and elaboration of texts on Korean high school students' reading comprehension. Her results showed that elaboration is more facilitative than simplification. Simplification also facilitated the participants' reading comprehension; however, low-proficiency students did not significantly benefit from simplification.

As the empirical studies cited above indicate, researchers have separately addressed the effects of language proficiency, content schemata, and linguistic simplification on L2 reading comprehension. Some researchers investigated the interactive effects of the combination of the two variables we focused on, namely, simplification and L2 proficiency. However, the literature provides no consistent evidence as to (a) whether sheer syntactic or sheer lexical simplification or their combination will have the same interactive effect as content schemata on reading comprehension, (b) whether this language-content interaction is the same for higher and lower EFL proficiency levels, and (c) whether higher-proficiency-level students perform in the same way on content-familiar and content-unfamiliar texts as lower-proficiency-level students. The above questions provided the impetus for the present study, which is mainly concerned with the interactive effects of the above-mentioned independent variables on EFL reading comprehension and recall. Therefore, the following research questions were posed:

1. Do content schemata have any significant effect on the relationship among EFL proficiency, reading comprehension, and recall?
2. Are there any significant interactions among the effects of content schemata, EFL proficiency, and lexical or syntactic simplification on EFL reading comprehension and recall?

Method

Participants

The participants were 240 male Iranian EFL students between the ages of 20 and 46, with an average age of 30. They were learning English at a private English institute (Kish Language School) in Tehran. The participants were divided into four groups of 60 students (30 of a low- and 30 of a high-proficiency level). The subjects' scores on the Nelson English Language Test (NELT; Fowler & Norman, 1976) were used as a consistent criterion for assigning the participants into the proficiency levels. Those receiving scores below the mean were categorized as low-proficiency students, and those obtaining scores above the mean as high-proficiency students. The results of two separate one-way ANOVAs confirmed the homogeneity of the subgroups within the high- and low-proficiency levels.

Instrumentation

Two types of instruments were used in this study. The first type included an NELT and eight reading comprehension tests that we constructed. NELT was chosen based on a pilot study with a

representative sample of participants similar to the research participants described above. A data analysis confirmed that the test was appropriate for determining the students' general English proficiency levels. The second type of instrument included eight reading comprehension tests based on two texts. One of the texts, an extract from the biography of the Prophet Muhammad (P.B.U.H), had content related to Islam. The other text was an extract from the biography of Joseph Smith, a non-Muslim religious figure. The content of the first text is supposed to be much more familiar to the Muslim participants than that of the second one. This is the very reason for choosing these texts. The texts were similar as far as genre, length, and linguistic difficulty are concerned. They were both extracts from the biographies of religious characters. They consisted of the same number of T-units. Using the Fog index of readability, the linguistic difficulty levels of the content-familiar and the content-unfamiliar texts were calculated to be 33.7 and 37.2. Further, a sample group of 42 participants, in a pilot study, were asked to identify the unknown words and the difficult structures of both of the texts. The result showed that the texts had almost the same number of unknown words and difficult structures. Thus, all the text variables except the content were controlled. Each of these texts appeared in four versions: original, lexically simplified, syntactically simplified, and lexically-syntactically simplified. Altogether, eight reading comprehension tests with the following text types were used:

1. Content-familiar, original text (CF.O)
2. Content-familiar, syntactically simplified text (CF.SS)
3. Content-familiar, lexically simplified text (CF.LS)
4. Content-familiar, syntactically and lexically simplified text (CF.SLS)
5. Content-unfamiliar, original text (CU.O)
6. Content-unfamiliar, syntactically simplified text (CU.SS)
7. Content-unfamiliar, lexically simplified text (CU. LS)
8. Content-unfamiliar, syntactically and lexically simplified text (CU.SLS)

For each of the texts, 14 MC test items were developed. The tests were piloted and pre-tested with a sample group of participants.

Note that one purpose of the present study was to investigate the effect of content schemata on reading comprehension. Hence, the items with high or low facility indices were not discarded, except when an item was realized to be too easy due to the directness of its answer. When the data analysis and item modification were accomplished, ten MC items were chosen for each of the texts.

Concerning the psychometric properties of the reading comprehension tests, the reliabilities of the content-familiar and the content-unfamiliar tests, as estimated using K-R 21, were .70 and .63. Both of the tests were validated against the reading comprehension sub-test of a TOEFL. The content-familiar test had a validity index of .77, and the content-unfamiliar one had a validity of .69.

To simplify the unmodified texts, we asked a pilot group of participants to read the texts and underlined the sentences whose grammatical structures were difficult for them, and they made a separate list of the unknown words for each of the texts. The lexical and syntactic simplification procedures are elaborated on below.

Lexical and Syntactic Simplification

To simplify the texts lexically, we replaced the unknown words with synonyms wherever possible. When an unknown word did not lend itself to being substituted with a one-word synonym, its definition was given in bold print between parentheses immediately after it. The definitions were given in bold print so that the participants' attention could be drawn to them. The synonyms and definitions were taken from the 1992 edition of *Longman Dictionary of Contemporary English*.

As stated in this dictionary, all the definitions in it were written using a list of defining vocabulary items. The defining words were carefully chosen after a thorough study of all the well-known frequency lists of English words. In this dictionary, if a definition includes a word that is not in the list, it is written in small capital letters. In such cases, some experienced EFL teachers were invited to choose as easy a definition or synonym for the given unknown word as possible. Although frequency lists are not considered valid predictors of text comprehension level by ESL and EFL readers, the present research confirms the overall finding that "better readers tend to have larger sight vocabularies, and those with larger sight vocabularies tend to be better readers" (Pulido, 2004, p. 473). A two-way relationship between lexical knowledge and reading comprehension of L2 readers has been documented. Therefore, we assume that frequency standards may be considered a consistent criterion for the simplicity level of the lexical definitions in the present study.

The next step was syntactic simplification. The sentences considered difficult by the participants were broken down into simple sentences. For the sake of consistency, we stuck to Richards, Platt, and Platt's (1992) definition of a simple sentence, "a sentence which contains only one predicate" (p. 70). The sentences the learners identified as difficult were some compound (sentences with more than one predicate) and complex sentences (especially sentences with adjective clauses).

Test Administration Procedure

To avoid participant fatigue, the content-familiar and the content-unfamiliar tests were administered in two separate sessions with an interval of four days. The participants were asked to read the texts and answer the MC items. Then, the participants were asked to read the texts once more and write down what they could remember on their recall answer-sheets without looking back at the text and the MC questions. Here, our main concern was to measure the participants' recall of what they read rather than their EFL writing ability. Accordingly, to avoid an adverse effect of low EFL writing ability on the participants' recalls, they were asked to write their recalls in their first language, Persian. The time allocated to each test was 45 minutes, which was determined to be suitable in the pilot study. The eight reading comprehension tests were randomly assigned to the four groups.

Scoring Procedure

The recall protocols were scored by counting the correct propositions written by the participants.

For the sake of a consistent scoring procedure, we defined propositions as the relationships between a predicate and its arguments. According to Brown and Yule (1983), this is the most frequently used definition in the text analysis literature.

After a short panel discussion, the raters agreed upon the above definition of a proposition. Then, the protocols were scored. Thirty of the recall protocols, which were randomly selected, were scored by one of the researchers and two experienced EFL teachers. The correlation indices between the scores given to the 30 recall protocols by the raters ranged from .79 to .85.

Results

The data consisted of the participants' answers to the MC comprehension questions and the analysis of the written protocols. As mentioned above, the recall protocols were analyzed for the quantity of propositions remembered from the text. The data were fed to an SPSS database for statistical analysis. The descriptive statistics in Tables 1 and 2 provide an overall view of the participants' performance on the reading comprehension and recall tests.

Table 1. *Descriptive statistics of MC test scores*

Text version	<i>N</i>	Minimum	Maximum	<i>M</i>	<i>SD</i>
CF.O	59	3.00	10.00	7.44	2.05
CF.SS	58	3.00	10.00	6.88	1.84
CF.LS	59	3.00	10.00	7.21	1.76
CF.SLS	57	2.00	10.00	7.00	1.99
CUF.O	59	1.00	8.00	4.39	1.71
CUF.SS	58	1.00	8.00	4.76	1.81
CUF.LS	59	2.00	9.00	5.41	1.52
CUF.SLS	57	1.00	8.00	5.14	1.82

Table 2. *Descriptive statistics of recall scores*

Text version	<i>N</i>	Minimum	Maximum	<i>M</i>	<i>SD</i>
CF.O	59	16.4	54.72	33.97	9.03
CF.SS	58	11.32	62.26	34.31	10.36
CF.LS	59	11.32	56.60	29.14	9.82
CF.SLS	57	9.43	45.28	26.92	9.10
CUF.O	59	2.25	44.79	15.56	7.66
CUF.SS	58	3.12	37.50	16.20	8.06
CUF.LS	59	3.12	39.58	19.19	7.97
CUF.SLS	57	3.12	41.68	18.26	9.29

To examine the main effects of the independent variables on the participants' reading comprehension and recall and their interaction effects, we ran two $2 \times 4 \times 2$ (Content \times Linguistic Simplification \times Proficiency) analyses of variance (ANOVA) with repeated measures: one for the MC reading comprehension test scores and one for the recall scores. In a repeated measures analysis, because each participant had taken one content-familiar and one content-unfamiliar test, the variable of content was treated as the within-participants factor, and the variables of linguistic simplification and EFL proficiency as between-participant factors. The results of the repeated measures analysis related to the MC test scores are presented in Tables 3

and 4, and the recall scores, in Tables 5 and 6.

Table 3. *Tests of within-participants contrast in MC tests*

Source	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
CON	1	566.228	307.874	.000
CON × LAN	3	9.616	5.228	.002
CON × PRO	1	2.789	1.517	.219
CON × LAN × PRO	3	2.019	1.098	.351

Note. CON = content schemata; LAN = linguistic simplification; PRO = EFL proficiency.

Table 4. *Tests of between-participants effects in MC tests*

Source	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
LAN	3	4.220	1.024	.383
PRO	1	128.420	31.171	.000
LAN × PRO	3	2.272	0.55	.648

Note. CON = content schemata; LAN = linguistic simplification; PRO = EFL proficiency.

Table 5. *Tests of within-participants contrast in recall scores*

Source	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
CON	1	22292.008	746.710	.000
CON × LAN	3	780.164	26.133	.000
CON × PRO	1	7.428	0.249	.618
CON × LAN × PRO	3	4.249	0.142	.934

Note. CON = content schemata; LAN = linguistic simplification; PRO = EFL proficiency.

Table 6. *Tests of between-participants effects in recall scores*

Source	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
LAN	3	143.569	1.329	.266
PRO	1	5629.361	52.124	.000
LAN × PRO	3	60.150	0.557	.644

Note. CON = content schemata; LAN = linguistic simplification; PRO = EFL proficiency.

As illustrated in Tables 3–6, the variables of content schemata (the within-participants factor) and EFL proficiency (a between-participants factor) had significant effects on reading comprehension and recall, whereas the linguistic simplification (a between-participants factor) was found to have no effect on reading comprehension and recall at a .05 significance level. In reading comprehension scores, the *F* ratio of the content-schemata effect was 307.874, while that of the linguistic simplification effect was 1.024. In the recall scores, the *F* ratio of the content-schemata effect (746.710) was much greater than the *F* ratio of the linguistic simplification (1.329). Therefore, we conclude that the content schemata had a greater effect on both reading comprehension and recall than the linguistic simplification (i.e., lexical and/or syntactic simplification).

To test the first null hypothesis, we calculated the correlation between EFL proficiency and reading comprehension on both the content-familiar and content-unfamiliar tests. These correlations were significant in the case of both reading comprehension and recall scores. Table 7 presents the correlations of EFL proficiency with the MC reading comprehension test scores and recall scores on the content-familiar and content-unfamiliar tests.

As illustrated in Table 7, EFL proficiency had a significant correlation with both MC reading comprehension test scores and recall scores. This significant correlation existed for both content-familiar and content-unfamiliar tests. In other words, the content did not affect the relationship between EFL proficiency and reading comprehension or recall. These results reflect the lack of content and proficiency interaction. Therefore, the null hypothesis stating that content schemata have no effect on the relationship between EFL proficiency and EFL reading comprehension and recall (i.e., the first null hypothesis) was not rejected.

Table 7. *Pearson correlation of M-C tests scores and recall scores with EFL proficiency (N=59)*

	PRO	CF.O	CUF.O
MC tests (N = 59)			
PRO	1.000	.581**	.291*
Recall (N = 59)			
PRO	1.000	.470**	.603**

Note. N = number of subjects; PRO = EFL proficiency; CF.O = content-familiar original test; CUF.O = content-unfamiliar original test.

* $p < .05$, ** $p < .01$.

Among the three variables of content, proficiency, and linguistic simplification, the effects of content and linguistic simplification interacted. This interaction has been shown in Tables 3 and 5. Therefore, the null hypothesis of no interaction among the effects of content schemata, linguistic simplification, and EFL proficiency on reading comprehension and recall (i.e., the second null hypothesis) was rejected. To locate the differences among the groups, we ran Tucky Tests. The results for the MC reading comprehension test scores and recall scores are shown in Tables 8 and 9.

Table 8. *Multiple comparisons of content \times linguistic simplification MC test scores of groups*

M	Groups	G ₁	G ₂	G ₃	G ₄	G ₅	G ₆	G ₇	G ₈
7.44	G ₁ CF.O		0.56	0.23	0.44	3.05*	2.67*	2.03*	2.30*
6.88	G ₂ CF.SS			0.33	0.12	2.49*	2.10*	1.47*	1.74*
7.21	G ₃ CF.LS				0.21	2.82*	2.44*	1.80*	2.08*
7.00	G ₄ CF.SLS					2.61*	2.22*	1.59*	1.86*
4.39	G ₅ CUF.O						0.39	1.02*	0.75
4.76	G ₆ CUF.SS							0.63	0.36
5.41	G ₇ CUF.LS								0.27
5.14	G ₈ CUF.SLS								

Note. * $p < .05$.

As shown in Tables 8 and 9, all the content-familiar groups outperformed the content-unfamiliar ones in both the MC reading comprehension test scores and the recall scores, showing that the content contributed more to both reading comprehension and recall than did linguistic simplification.

The point of concern here is the interactions between the variables of content and linguistic simplification. As shown in Tables 4 and 6, linguistic simplification had no major effect. However, its effect must be interpreted in the light of the interaction between the content and

linguistic simplification (see Tables 3 and 5).

Table 9. Multiple comparisons of content \times linguistic simplification recall scores of groups

<i>M</i>	Groups	G ₁	G ₂	G ₃	G ₄	G ₅	G ₆	G ₇	G ₈
33.97	G ₁ CF.O		0.34	4.32	7.05*	18.41*	17.77*	14.54*	15.71*
34.31	G ₂ CF.SS			4.66	7.39*	18.74*	18.11*	14.88*	16.05*
29.14	G ₃ CF.LS				2.73	14.08*	13.46*	10.21*	11.39*
26.92	G ₄ CF.SLS					11.36*	10.72*	7.49*	8.66*
15.56	G ₅ CUF.O						0.63	3.87	2.70
16.20	G ₆ CUF.SS							3.23	2.06
19.19	G ₇ CUF.LS								1.17
18.26	G ₈ CUF.SLS								

Note. * $p < .05$.

As for the interactions related to the MC reading comprehension tests, lexical simplification had a significant effect on the reading comprehension of the content-unfamiliar text; however, it had no significant effect on the participants' comprehension of the content-familiar text. Figure 1 gives a graphic representation of the group means related to the MC test scores.

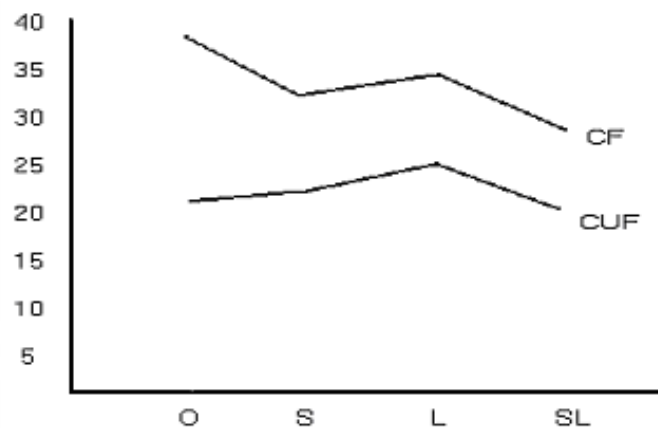


Figure 1. Content \times linguistic simplification interaction in MC tests.

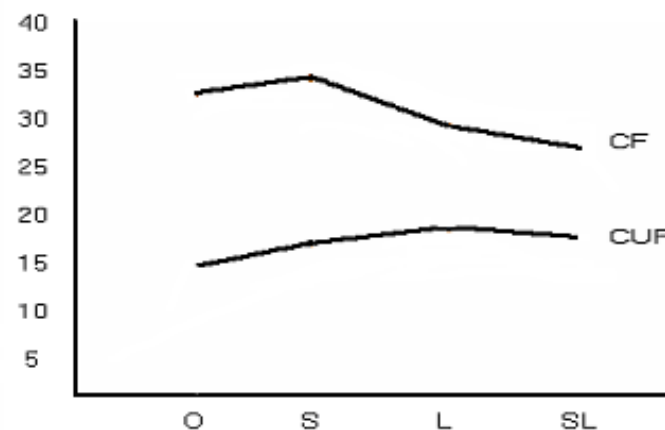


Figure 2. Content \times linguistic simplification interaction in recall scores.

Concerning interactions with the recall scores, as illustrated in Table 9, the CF.O and CF.SS groups significantly outperformed the CF.SLS group. That is to say, the combination of the syntactic and lexical simplification significantly impeded the participants' recall of the content-familiar text, while it facilitated the recall of the content-unfamiliar text, although not significantly. The means of the groups related to recall scores and interaction between the content and linguistic simplification are presented in Figure 2.

Discussion

We found content schemata to have a greater effect than lexical or syntactic simplification on both EFL reading comprehension and recall. This is compatible with the results of studies by Carrell (1987), Johnson (1982), Kang (1992), Oh (2001), Steffensen and Joag-Dev (1984), and Steffensen, Joag-Dev, and Anderson (1979). Of course, these studies used different kinds of texts.

Notably, as the results of the repeated measures analyses (Tables 4–6) show, language proficiency had a significant effect on both reading comprehension and recall, whereas linguistic simplification (syntactic or lexical) showed no significant effect. Contrary to our expectation, language proficiency and linguistic simplification did not affect reading comprehension and recall in the same or similar ways. As a reader variable, language proficiency was expected to enhance the participants' comprehension and recall of the texts, as it did in the present study. As a text variable, linguistic simplification was expected to be a mirror image of language proficiency and to have similar effects on reading comprehension and recall. However, the linguistic simplification did not show any significant effect on reading comprehension and recall. The same result was found by Floyd and Carrell (1987). In their study, the syntactic complexity of the text did not significantly affect comprehension.

As Widdowson (1979) maintains, by simplification, we mean bringing the language of the original within the scope of the learner's transitional linguistic competence. However, as Widdowson (1978) and Davies (1984) argue convincingly, syntactic simplification might result in distorting the authenticity value of texts and eventually render texts less readable. In the present study, the linguistic simplification did not significantly enhance the reading comprehension of the EFL learners. The inhibiting effect of the lexical simplification and the non-significant positive effect of the syntactic simplification on reading comprehension in the above-mentioned studies and the present one may cast serious doubts on the validity of the readability formulae as sound predictors of text difficulty. This also means that estimates of text difficulty should go beyond measures of linguistic complexity and that reading comprehension should be reformulated as levels of understanding based on the interaction among text variables and reader factors. Along the same lines, Fulcher (1997), stating that readability formulae are just functions of sentence length and vocabulary, criticizes the formulae for not taking into account the other important text and reader variables in reading. As stated above, in this study, content showed a greater effect than linguistic simplification on both reading comprehension and recall. Therefore, content is a greater determining factor than language in text difficulty. This is in line with the findings of McAdams (1993), which showed that topic and reader interest were

more important than sentence length. Also, the interaction among content, linguistic simplification, and reading comprehension, as documented in this study, suggests that in text-readability formulae, both the language and content should be taken into account. That is to say that different readability formulae may be needed for content-familiar and content-unfamiliar texts.

As stated above, language proficiency showed a significant positive correlation with both reading comprehension and recall regardless of content. This goes against the finding of Carrell (1983). She found that participants with advanced L2 proficiency recalled more from content-unfamiliar texts than from the content-familiar ones; however, the converse was true of intermediate-proficiency participants. These participants received higher scores in the recall of content-familiar than content-unfamiliar texts.

The findings of the present study are compatible with those of Carrell (1983), both indicating the existence of a language proficiency threshold above which content and proficiency interaction appears. Lee (1986) found that his participants could recall more from a content-unfamiliar text than a content-familiar one. This goes against our results. In the present study, the recall from the content-familiar texts was significantly higher than the recall from the content-unfamiliar texts. Considering the content by proficiency interaction in Carrell's (1983) study, the differing proficiency levels are a plausible explanation for the disagreement between our findings and those of Lee (1986).

Studies by Koh (1985) and Peretz and Shoham (1990) also showed contrary results: the highest performance on a given text often was not obtained by the group that was expected to be favored by the text. Accordingly, students do not necessarily do better on materials in their own academic fields. Alderson (2000) stated that this may be due to a superior linguistic proficiency compensating for ignorance of the subject matter.

Similarly, Clapham (1996) studied the effect of content, specifically, subject matter knowledge, and the relationship between the language ability of students taking the IELTS test of reading for academic purposes and their ability to understand text in and out of their subject discipline. She found two linguistic thresholds. The first one, at a score of roughly 60% on her grammar test, represented a level of linguistic knowledge below which students were unable to understand texts even in their own subject disciplines. The second, at a score of roughly 80% on the same test, represented a level of linguistic knowledge above which the participants had little difficulty reading texts outside their own disciplines. The crucial area, in which subject knowledge could facilitate understanding of texts within one's own subject area, was 60–80% on the test.

As stated above, we found that the content and linguistic simplification interacted. The lexical simplification had a significantly facilitative effect on reading comprehension of the content-unfamiliar texts, but it had an impeding effect on the participants' comprehension of the content-familiar text, although the effect was not significant. This may indicate that when the content is familiar, readers can guess the meanings of unknown words, and linguistic simplification does not improve the readers' comprehension. However, when the content is unfamiliar, readers cannot guess the meanings of new words. Hence, new lexical items will add to the readers' problems, and lexical simplification will enhance their comprehension.

Note that the impeding effect of the lexical simplification on the participants' comprehension of the content-familiar texts was not significant. As mentioned above, in the process of the lexical simplification, some unknown words did not lend themselves to being substituted with synonyms. In such cases, the definitions of the words were given next to them in parentheses. The hindering effect of the lexical simplification on the comprehension of the content-familiar text may be attributed to the parenthetical definitions of the words. That is to say that these definitions distracted the participants and decreased their comprehension. It is difficult to explain why the same procedure of lexical simplification did not impede the comprehension of the content-unfamiliar text but facilitated it significantly. One possible interpretation is that the readers may have used a conceptual rather than a syntactic strategy requiring more vocabulary, content, and world knowledge.

Conclusion

Our findings show that content schemata have a greater effect than linguistic simplification on both EFL reading comprehension and recall. This suggests that in EFL reading, content is of the utmost importance. Therefore, we recommend that material developers not choose content-unfamiliar texts to present new linguistic items to EFL readers.

The participants in this research tended to understand and recall the simplified version of the content-unfamiliar text better than its original version, while this was not true of the content-familiar text. Therefore, EFL teachers and students should remember that EFL students may have problems reading original texts for which they do not have the required background knowledge.

The interaction between the content and linguistic simplification indicates that the readability of a text should be considered as the interaction of a multitude of factors. The results show that the linguistic simplification facilitated the comprehension and recall of the content-unfamiliar text, whereas it had an impeding effect on the comprehension and recall of the content-familiar text. This finding, along with the previous literature on text selection, may shed new light on more valid approaches to reading materials for EFL students with different proficiency levels and schematic backgrounds. The findings confirm the previous research indicating that input modification (simplification in this case) does not necessarily result in greater readability for EFL learners. However, the findings of this study need to be verified against future research involving more representative samples comprising both males and females. Similarly, the small size of the reading test batteries might have distorted the pattern of findings. Therefore, further research needs to be conducted before solid generalizations can be made.

References

- Alderson, J. C. (1984). Reading in a foreign language: A reading problem or a language problem? In J. C. Alderson & A. H. Urquhart (Eds.), *Reading in a Foreign Language* (pp. 1–27). London: Longman.

- Alderson, J. C. (2000). *Assessing reading*. Cambridge: Cambridge University Press.
- Ammon, M. S. (1987). Patterns of performance among bilingual children who score low in reading. In S. R. Goldman & H. T. Trueba (Eds.), *Becoming literate in English as a second language* (pp. 71–105). Norwood, NJ: Ablex.
- Bartlett, F. C. (1932). *Remembering*. Cambridge: Cambridge University Press.
- Blau, E. K. (1982). The effect of syntax on readability for ESL students in Puerto Rico. *TESOL Quarterly*, 16, 517–28.
- Britton, B. K., Gulgoz, S., & Glynn, S. (1993). Impact of good and poor writing on learners: Research and theory. In B. K. Britton, A. Woodward, & M. Binkley (Eds.), *Learning from textbooks: Theory and practice* (pp.1–46). Hillsdale, NJ: Lawrence Erlbaum Publishers.
- Brown, G., & Yule, G. (1983). *Discourse analysis*. Cambridge: Cambridge University Press.
- Byrd, H. P. (2000). *It's all the same grammar: Re-thinking grammar at various proficiency levels*. Retrieved from <http://www.gsu.edu/~eslhpb/grammar/info/same.htm#Tradition>
- Carrell, P. L. (1987). Content and formal schemata in ESL reading comprehension. *TESOL Quarterly*, 21, 461–481.
- Carrell, P. L. (1984). Schema theory and ESL reading: Classroom implications and applications. *Modern Language Journal*, 68, 332–343.
- Carrell, P. L. (1983). Three components of background knowledge in reading comprehension. *Language Learning*, 33, 183–207.
- Clapham, C. (1996). *The development of IELTS: A study of the effect of background knowledge on reading comprehension*. Cambridge: Cambridge University Press.
- Davies, A. (1984). Simple, simplified and simplification. In J. C. Alderson & A. H. Urquhart (Eds.), *Reading in a Foreign Language* (pp. 181–195). London: Longman.
- Floyd, P., & Carrell, L. P. (1987). Effects on ESL reading of teaching cultural content schemata. *Language Learning*, 37, 89–108.
- Fowler, W. S., & Norman C. (1976). *Nelson English Language Tests*. Canada: Thomas Nelson and Sons Ltd.
- Fulcher, G. (1997). Text difficulty and accessibility: Readability formula and expert judgment. *System*, 25, 479–513.
- Grabe, W. (1997). *Reading research and its implications for reading assessment* (Language Teaching Resource Center Paper). Flagstaff: Northern Arizona University.
- Johnson, P. (1982). Effects on reading comprehension of building background knowledge. *TESOL Quarterly*, 16, 503–516.
- Kang, H. W. (1992). The effects of culture-specific knowledge upon ESL reading comprehension. *School of Education Review*, 4, 93–105.
- Koh, M. Y. (1985). The role of prior knowledge on reading comprehension. *Reading in a Foreign Language*, 3, 375–380.
- Langer, J. A., Bartolome, L., Vasquez, O., & Lucas, T. (1990). Meaning construction in school literacy tasks: A study of bilingual students. *American Educational Research Journal*, 27, 427–471.
- Lee, J. F. (1986). Background knowledge and L2 reading. *Modern Language Journal*, 70, 350–354.
- McAdams, K. C. (1993). Readability reconsidered: A study of reader reactions to Fog indexes. *Newspaper Research Journal*, 14, 50–59.
- Mountford, A. (1975). Discourse analysis and the simplification of reading materials for ESP.

- Unpublished MLitt thesis, University of Edinburgh, Edinburgh, U.K.
- Oh, S. Y. (2001). Two types of input modification and EFL reading comprehension: Simplification versus elaboration. *TESOL Quarterly*, 35, 69–96.
- Parker, K., & Chaudron, C. (1987). The effects of linguistic simplification and elaborative modifications on L2 comprehension. *University of Hawaii Working Papers in ESL*, 6, 107–133.
- Peretz, A. S., & Shoham, M. (1990). Testing reading comprehension in LSP: Does topic familiarity affect assessed difficulty and actual performance? *Reading in a Foreign Language Journal*, 7, 447–454.
- Pulido, D. (2004). The relationship between text comprehension and second language incidental vocabulary acquisition: A matter of topic familiarity. *Language Learning*, 3, 469–523.
- Richards, J. C., Platt, J., & Platt, H. (1992). *Longman dictionary of language teaching and applied linguistics*. London: Longman Group Limited.
- Steffensen, M. S., Joag-Dev, C., & Anderson, R. C. (1979). A cross-cultural perspective on reading comprehension. *Reading Research Quarterly*, 15, 10–29.
- Steffensen, M. S., & Joag-Dev, C. (1984). Cultural knowledge and reading. In J. C. Alderson & A. H. Urquhart (Eds.), *Reading in a Foreign Language* (pp. 48–61). New York: Longman.
- Strother, J. B., & Ulijn, J. M. (1987). Does syntactic rewriting affect English for science and technology (EST) text comprehension? In J. Devine, P. L. Carrell & D. E. Eskey (Eds.), *Research in reading in English as a second language*. Washington, D.C.: TESOL.
- Widdowson, H. G. (1978). *Teaching language as communication*. Oxford: Oxford University Press.
- Widdowson, H. G. (1979). *Explorations in applied linguistics*. Oxford: Oxford University Press.
- Yano, Y., Kong, M., & Ross, S. (1994). The effects of simplified and elaborated texts on foreign language reading comprehension. *Language Learning*, 44, 189–219.

About the Authors

Mohammad Hossein Keshavarz is a professor of applied linguistics at Tarbiat Moallem University of Tehran, Iran. He is the founding editor of the *Iranian Journal of Applied Linguistics* and has published on linguistic and TEFL issues in a number of national and international journals. He is also the author of five English textbooks. His research interests include sociolinguistics, phonetics and phonology, reading, error analysis, and discourse analysis.

Mahmood Reza Atai is an assistant professor at Tarbiat Moallem University of Tehran, Iran. His research areas include reading comprehension, vocabulary learning, ESP/EAP materials production, and program evaluation. He has published three books and 12 articles in national and international journals, and he has presented 15 papers at national and international conferences.

Hossein Ahmadi holds an M. A. degree in TEFL from Tarbiat Moallem University and has taught EFL in different English institutes in Tehran. He is currently an EFL instructor at Azad University in Malayer, Iran.