

READABILITY OF ESP TEXTBOOKS IN IRAN: A NEGLECTED ISSUE OR A TAKEN-FOR-GRANTED ONE?

Hassan Mohebbi¹, Akram Nayernia², Majid Nemati³, Behzad Mohebbi⁴

¹Faculty of Foreign Languages and Literature, University of Tehran, Iran

²Iran University of Science and Technology, Iran

³Faculty of Foreign Languages and Literature, University of Tehran, Iran

⁴University of Mohaghegh Ardabili, Iran

Phone: +984533749317, E-Mail: hassanmohebbi@ut.ac.ir

Abstract. *This study investigates the readability index of the ESP textbooks taught at Iranian universities. To this end, 51 randomly selected texts from 11 ESP textbooks published by The Organization for Researching and Composing University Textbooks in the Humanities (SAMT) were analyzed. Seven main readability indexes, including Flesch Reading Ease Score, Gunning Fog, Flesch-Kincaid Grade Level, The Coleman-Liau Index, The SMOG Index, Automated Readability Index, and Linsear Write Formula were used to assess the readability of the texts. Surprisingly enough, the data analysis revealed that the majority of the texts were difficult, fairly difficult, and very difficult. The findings underscore that the texts used in the ESP textbooks at Iranian universities are not appropriate for ESP students in terms of the readability of the texts. It is concluded that readability of ESP books is a neglected issue that needs more in-depth consideration. The findings of the study might be useful for ESP textbooks' authors and materials developers. The pedagogical implications of the study and further research directions are explained in detail.*

Key words: *ESP, ESP Textbooks, Readability Indexes, SAMT, Iran*

1. INTRODUCTION

In the last three decades, there have been a good number of studies investigating the different aspects of English for Specific Purposes (ESP) (for review see Belcher, 2004; Brunfaut, 2014; *Grosse & Voght*, 2012; Johns & Dudley-Evans, 1991; Swales, 2000). To date, studies have mostly focused on needs analysis (e.g., Belcher, 2006), genre analysis (e.g., Skulstad, 1999), technology applications (e.g., Arnó Macià, 2012), curriculum and materials development (e.g., Basturkmen, 2003; McDonough, 2010), and testing issues in ESP (e.g., Davies, 2001; Emery, 2014; O'Sullivan, 2012; Skehan, 1984). However, there is a gap regarding textbook evaluation, in particular the readability of the texts included in ESP books. To our best knowledge, to date, hardly any research has been done to investigate the readability of ESP books. To bridge this gap in the ESP literature, this study examines the readability of ESP texts used in the ESP books taught at Iranian universities in undergraduate and graduate programs. The following section of the article reviews briefly the status of ESP in Iran.

2. ESP IN IRAN

According to the Curriculum and Textbooks Development Office (www.talif.sch.ir) which is the official organization for educational research and planning working under supervision of Ministry of Education, Iranian students start to learn English when they enter middle and high school levels. They have English course for six years. Having a highly centralized English language teaching curriculum, the local teachers have no voice and place in policymaking and expressing their viewpoints on the quantity and quality of the curriculum. Students have English classes between two and four hours each week. But, the output of English classes is not satisfactory due to various factors such as the curriculum, the textbooks, and the teachers. The main focus of English classes is on grammar, vocabulary, and reading comprehension; listening is not included in the syllabus and there are few speaking drills which are mainly aimed to practice grammar (Sadeghi & Richards, 2016).

According to higher education system curriculum in Iran, university students need to pass a two-credit Basic English course, a three-credit General English course, and a three-credit ESP course; however, based on their majors, these credits can be more or less in some cases. As Soodmand Afshar and Movassagh (2016) and Hayati (2008) rightly asserted, the majority of Iranian students and teachers are not satisfied with the ESP and EAP (English for Academic Purposes) courses. The reason behind this dissatisfaction might be the philosophy and practice of ESP in the context of Iranian universities. The ESP courses in Iran are merely designed to enable students to read for understanding texts and do a multiple-choice reading comprehension test and in some cases translate the texts just for the sake of doing an assignment (Hayati, 2008) because English in Iran is considered as a tool to have access to new knowledge and technology (Sadeghi & Richards, 2016) and obtain industrial and economic self-sufficiency (Atai & Mazlum, 2013) and there is a little oral interaction out of English classes especially in the last decade due to the sanctions and political issues surrounding Iran.

Hayati goes as far as to claim that there are tri-partite problems regarding ESP course in Iran, namely the teacher, the time, and the textbooks. In fact, there is lack of teachers who have expertise in both language and content (Soodmand Afshar & Movassagh, 2016). More importantly, there are no teacher training and teacher education programs devoted specifically to ESP teachers to educate teachers who have mastery of English and the technical knowledge. Moreover, the time devoted to English, in general, and ESP, in particular, is so limited that ESP courses do not yield any positive result. Finally, yet importantly, the textbooks compiled and used are not in line with the current theories of language education and ESP. The texts and tasks used in these ESP books are not communicative-oriented and there is no place for authentic tasks which need students' genuine interaction; most importantly, the texts used in the ESP textbooks in Iran seem beyond the English proficiency level of Iranian Students. This situation does not satisfy the students' needs in academic and vocational contexts; however, unfortunately, there is seemingly no will to take steps to enhance the quality of ESP courses in Iran. Due to the fact that textbooks are "primary learning tool given to students" (Sawyer, 1991, p. 307), investigating the textbooks is of primary importance.

We should add another factor for the failure of ESP courses in Iran, namely the general English language proficiency of ESP students. The ESP students are expected to be at intermediate and upper-intermediate level of general English language proficiencies;

but the evidence in Iran, based on our and our colleagues' experiences in ESP classes across the country, indicates that although the situation has been improved in recent years, the majority of ESP students are at the beginner or pre-intermediate level of English language proficiency. In addition, this situation varies from university to university and students' field of study plays a key role in their language proficiency. You may study Soodmand Afshar and Movassagh (2016) for a recent critical analysis of EAP education in Iran.

The Organization for Researching and Composing University Textbooks in the Humanities (SAMT), founded in 1985, is affiliated to the Ministry of Science, Research, and Technology of Iran. One of the main missions of SAMT is to publish textbooks in the humanities needed by the Iranian universities at the academic level, in particular ESP textbooks. To date, SAMT has published more than 150 ESP books for different majors, including science, humanities, arts, medicine, and engineering. However, they suffer from the one-size-fits-all philosophy and, as already highlighted, are not in line with the current theories of language learning and ESP. It should be mentioned that although SAMT has recently updated and has provided sample lesson and ESP textbook templates (www.samt.ac.ir), it does not seem to meet the needs of ESP students in Iranian context. Hence, the situation of ESP courses in Iran, in particular the ESP textbooks warrants further research to encourage and convince policymakers and materials developers to make considerable changes in ESP curriculum and textbooks. In this study, as already mentioned, we aim to investigate the readability of the texts used in the ESP textbooks which is not researched yet in Iranian context. The next section of the paper elaborates on the reading and the factors influencing reading, in particular readability of texts.

3. READING: THE FACTORS INFLUENCING READING

Armbruster and Anderson (1988) point out that we should take into account the following variables in selecting a textbook: the readability level of texts, appropriate use of illustrations, using clear subtitles, and using appropriate connective words and phrases. Reading is a complex cognitive activity which involves simultaneous linguistic processing and activating prior knowledge, storing information, and monitoring comprehension (Pulido, 2007). Specific reader characteristics, text properties, and the context in which reading takes place affect reading (Kendeou, Muis, & Fulton, 2011). Text-related variables, including readability, word characteristics, text characteristics, the presence of contextual clues, and reader-related variables like second language proficiency, especially second language vocabulary and grammar knowledge, cognitive and mental effort of second language learners, and world/background knowledge influences second language reading; in a nutshell, reading is a cognitive process which involves the interaction of multilevel linguistic features such as word, semantics, syntax, and cohesion (Fox, 2009; Leroy, Kauchak, & Mouradi, 2013; Sung, Chen, Cha, Tseng, Chang, & Chang, 2015). White (2012) summarizes 34 text features which can be facilitator and/or inhibitor in reading, which increase the difficulty of a reading task, and are inherent in reading difficulty.

Cognitive load theory (Sweller, 1994) distinguishes three sources of cognitive load: extraneous, intrinsic, and germane which affect learning. Extraneous cognitive load which is caused by presenting the material in a poorly designed layout or including redundant material is cognitive processing that does not contribute to learning (Mayer & Moreno, 2010; Moreno & Park, 2010). Instructional design and materials, in this case reading texts,

should reduce the extraneous cognitive load resulted from inappropriate instructional procedures. Reducing the extraneous cognitive load is likely to free working memory capacity and enhance the germane cognitive load and facilitate learning (Sweller, 2005). Demonstrations, animations, simulations, exploratory environments, applying various modes, and modalities can impose extraneous cognitive load (Kalyuga, 2009). Therefore, one of the issues in reducing extraneous cognitive load should be checking the readability of the texts used in textbooks, especially ESP textbooks.

4. READABILITY

Tamor (1981) presents three operationalizations of text difficulty, namely text-based (objective), performance-based (behavioral), and a combination of the two (subjective). Readability as a text-based approach considers difficulty as characteristics such as syntax and vocabulary which are inherent in the text itself. Readability research has been the first attempts to provide guidelines for instructional texts (Sawyer, 1991).

It is argued that readability should be considered in textbook selection and materials development decisions (McConnell, 1983; Spinks & Wells, 1993). Wall (1969) called readability as a neglected criterion in textbook selection. Therefore, quantitative readability indexes can be useful in objective assessment of text difficulty or complexity and resolve the perceived problems related to relying on subjective human judgment about difficulty, complexity, and readability of a text (Sung, Dyson, Chen, Lin, & Chang, 2015). Considering readability as an indicator of text difficulty and complexity is in line with the verbal efficiency theory which assumes that processing capacity is limited; consequently, processing less frequent and rare words takes longer time and energy which makes a text difficult and/or complex (Reed & Kershaw-Herrera, 2016).

Providing students with accessible and well-matched texts to reader abilities has always been challenging for educators and material developers. In fact, text readability or text difficulty and accessibility is a crucial but neglected issue in Applied Linguistics (Fulcher, 1997). The philosophy behind readability is to produce texts matched to the abilities and needs of learners (Wray & Janan, 2013).

Since 1920, more than 50 formulas have been developed to assess text difficulty based on lexical or semantic features and sentence or syntactic complexity (Crossley, Grrenfield, & McNamara, 2008). Readability formulas have been the dominant paradigm for assessing text readability (Williamson, 2008) which can serve as handy and effective tools for (less-experienced) teachers in adopting materials (Parker, Hasbrouck, & Weaver, 2001). As Bailin and Greeneafstein (2001) rightly underscore, "a readability formula can return a numerical score, giving the user the sense of knowing the precise level of difficulty of a text" (p. 286). It seems that teachers' awareness of readability concerns might enhance reading instruction (Kasule, 2011). Readability indexes can help teachers and learners know the estimated level of a text in advance and decide whether to read a text or not (Zamanian & Heydari, 2012). More interestingly, these indexes are freely available, user-friendly, and can be interpreted easily.

As already mentioned, there are many readability formulas to assess the readability of texts. Mostly, the readability formulas take into account one or more of text variables such as percentage of high frequency easy words, percentage of hard words, average number of words per sentence, average number of syllables per word, number of single syllable words, or number of multiple-syllables words (Begeny & Greene, 2014). The

philosophy behind these formulas is that texts that contain shorter sentences and words and more frequent words are more readable than texts with longer sentences and less frequent words (Benjamin, 2012).

The next section explains the most reliable and commonly used readability indexes. These indexes are based on multisyllabic words and sentence length that are argued to be the two strongest predictor variables of reading difficulty (Schneider, 2011).

Flesch Reading Ease Score

The Flesch Reading Ease Score is based on a 0-100 scale. A higher score indicates the text is easier to read and comprehend and a low score warns that the text is difficult to read and understand. According to the Flesch Reading Ease Score, the best text is expected to contain shorter sentences and words. The score between 60 and 70 is largely considered acceptable (“Flesch–Kincaid readability tests”, 2016). Table 1 represents the criteria for assessing the readability of a text based on Flesch Reading Ease Score.

Table 1 The Flesch Reading Ease Score

Readability Score	Text Difficulty
0-29	Very Confusing
30-49	Difficult
50-59	Fairly Difficult
60-69	Standard
70-79	Fairly Easy
80-89	Easy
90-100	Very Easy

Gunning Fog

The Gunning Fog index assesses the readability of English writing. Precisely speaking, it indicates the years of formal education a person should have to understand a text on a first reading. The ideal index is 7 or 8. Indexes above 12 are interpreted too hard for most of persons to read (“Gunning fog index”, 2016).

Flesch-Kincaid Grade Level

Flesch-Kincaid Grade Level is used to assess the difficulty of technical manuals. The score obtained corresponds with a grade level. For instance, a score of 12.5 would indicate a student in 12th grade or first year of college can read and understand the text (“Flesch–Kincaid readability tests”, 2016).

The Coleman-Liao Index

Generally, The Coleman-Liao Index gives a lower grade value than the other readability indexes. Based on USA grade level, the Coleman-Liao Index provides the grade level that readers need to comprehend a text. For example, if the output index is 10.6, then the text would be appropriate for a 10th or 11th grade student (“Coleman–Liao index”, 2016).

The SMOG Index

The SMOG Index estimates the years of education needed to understand a piece of writing. It is widely used in assessing consumer-oriented healthcare materials (“Smog”,

2016). The SMOG Index is based on 100% reader comprehension and rates text difficulty two grade levels higher than the other comparable readability measures (Schneider, 2011).

Automated Readability Index

The Automated Readability Index assesses the understandability of a text. Despite the other readability indexes, it relies on a factor of characters per word instead of the usual syllables per word. It is widely used on all different types of texts (“Automated readability index”, 2016).

Linsear Write Formula

Linsear Write Formula is readability metric for English texts; it is specifically developed for USA Air Force for assessing the technical manuals’ readability (“Linsear write”, 2016).

5. THE EMPIRICAL STUDIES

Heydari and Riazi (2012) compared the readability of the texts used in English as a foreign language context (EFL) evaluated by university EFL teachers and graduate students and calculated by Flesch readability. They observed that there was a significant difference between the teachers’ and students’ evaluations and the Flesch readability.

Cline (1972) investigated the readability of community college textbooks and the reading ability of the students using these books. He found that 11 of the 17 textbooks assessed were above the reading levels of 50 per cent of the students and 7 of the textbooks were above the reading levels of at least 75 per cent of the students. He stressed on the importance of readability in selecting textbooks.

In a more recent research, Robison, Roden, and Szabo (2015) found that the social studies textbooks were written above grade level reading based on readability indexes.

6. THE STUDY

This study set out to investigate the readability of the ESP textbooks published by SAMT which are used in Iranian universities. To this end, 11 ESP textbooks were randomly selected to cover almost all majors, namely humanities, science, medicine, and engineering. The selected books were: English for the students of physics, English for the students of engineering, English for the students of visual arts, English for the students of mechanization and mechanics of agricultural machinery, English for the students of veterinary medicine, English for the students of sociology, English for the students of agricultural extension and education, English for the students of psychology, English for the students of sciences, English for the students of applied physics, and English for the students of medicine. Then, 51 texts from different chapters of the books were randomly selected and analyzed. In choosing the texts, we selected texts from the beginning, middle, and the last chapters of the books. Then, the texts were analyzed by three online readability software available at <http://www.readabilityformulas.com/flesch-grade-level-readability-formula.php>, read-able.com, and <https://readability-score.com>. The next section of the article explains the results of the data analysis in detail.

7. RESULTS

Grootens-Wiegers, De Vries, Vossen, and Van den Broek (2015) warned against that different readability formulas can yield significantly different results. Therefore, we used multiple readability formulas to supplement each other and have more reliable result and a much more comprehensive picture of the readability of the texts used in the ESP textbooks used in Iranian context.

Table 2 represents the readability indexes of the texts analyzed based on seven main readability indexes: Flesch Reading Ease Score, Gunning Fog, Flesch-Kincaid Grade

Table 2 The results of main readability indexes of the texts

Text	Flesch Reading Ease Score	Gunning Fog	Flesch-Kincaid Grade Level	The Coleman-Liau Index	The SMOG Index	Automated Readability Index	Linscar Write Formula
1	40.2	15.3	11.6	13	11	11.2	11.6
2	25.4	17.9	15	15	13.5	15.2	16
3	40.7	13.7	11.2	13	10.2	10.7	9.4
4	39.7	15.6	11.6	15	11.1	12.3	11.5
5	36.7	16.8	12.2	14	12	12.3	12.5
6	22.1	17.8	15	16	13.7	15.3	15.4
7	49.5	13.4	10.5	12	9.9	10.9	11.3
8	24.1	18.1	15.5	14	13.2	15.3	16.5
9	28.1	16.6	14.2	15	12.5	14.9	14.5
10	53.8	12.7	10.1	11	9.3	10.1	11.3
11	14.5	20.9	18.2	17	15.5	20.2	21.3
12	47.5	14.1	11.1	11	10.7	10.5	12.4
13	22.1	19.1	14.9	15	13.8	14.5	15.3
14	24.3	17.6	13.9	14	13.0	12.5	13.3
15	38.9	15.2	12.9	12	11.4	12.6	14.0
16	52.7	14.4	11.4	10	10.5	12.2	14.4
17	32.1	18.2	15.4	12	13.2	15.9	18.5
18	19.6	18.4	15.2	16	13.1	14.9	14.6
19	52.6	12.3	10.3	10	9.2	9.9	11.3
20	27.5	17.9	15.1	14	13	15.7	16.5
21	35.6	15.3	12.5	14	11.5	12.4	12.4
22	36.6	20.1	16.5	11	13.8	18	22.5
23	46.3	15.9	12.4	10	11.7	12.2	15.4
24	51.0	15.7	11.8	10	11.5	12.3	15.3
25	69.3	11.9	8.3	7.0	8.4	8.1	11.5
26	33.3	18.3	14.4	12	13.4	13.8	17.1
27	64.4	12.3	9.2	8.0	9.0	8.7	12.2
28	59.5	12.8	9.2	8.0	9.3	7.6	11.1
29	44.5	16.4	13	9.0	11.8	12.3	16.2
30	28.3	21.7	17	12	15.4	17.9	22.8
31	20.2	23.7	19.3	13	16.9	20.7	26.5
32	40.3	16	12.9	12	11.9	13.3	14.9
33	40.7	14.2	12.1	12	10.8	11.5	12.5
34	39.6	16.1	13.6	12	11.9	14	16.1
35	53.4	13.7	10.9	10	9.9	10.7	13.0
36	34.0	15.2	12.9	14	11.3	12.6	12.7
37	49.9	14.7	11.4	10	10.8	11	13.6
38	59.5	12.2	9.6	9.0	8.9	9.5	11.6
39	46.2	14.7	11.1	13	10.8	11.5	12.1
40	47.3	13.2	10.2	12	9.8	9.9	8.9
41	52.7	12.7	9.9	11	9.6	9.7	10.8
42	56.4	13.3	11	10	9.1	11.7	13.6
43	48.2	13.8	11.3	12	10.1	12.2	12.6
44	58.3	12.4	8.3	12	9	8.5	07.8
45	24.7	22.4	18.1	13	15.8	19.9	24.3
46	37.2	14.8	12.1	14	10.9	12.3	11.7
47	64.4	10.8	8.2	9.0	7.8	8.0	08.7
48	40.9	15.3	11.8	13	11.1	11.8	12.3
49	58.8	13.2	9.5	9.0	9.7	9.0	11.7
50	39.3	18.5	14.7	11	13.4	15.5	19.3
51	32.6	18.8	14.9	12	13.8	14.8	18.2

Level, The Coleman-Liau Index, The SMOG Index, Automated Readability Index, and Linsear Write Formula. As Table 2 clearly indicates, the majority of the texts were difficult to read and comprehend. It is imperative to mention that the data should also be interpreted contextually according to the ESP status in Iran. Therefore, the majority of the texts included in various ESP textbooks published by SAMT do not seem appropriate for Iranian ESP students because as already highlighted Iranian undergraduate and graduate students are at beginner and/or pre-intermediate English proficiency levels.

Table 3 shows the descriptive statistics of the main readability indexes of the texts analyzed.

Table 3. The descriptive statistics of the main readability indexes of the texts

Readability Index	Mean	SD	Variance
Flesch Reading Ease Score	41.28	13.30	177.01
Gunning Fog	15.80	2.90	8.45
Flesch-Kincaid Grade Level	12.61	2.65	7.06
The Coleman-Liau Index	12.01	2.24	5.06
The SMOG Index	11.54	2.05	4.22
Automated Readability Index	12.71	3.10	9.61
Linsear Write Formula	14.33	3.95	15.67

Figures 1 and 2 exhibit the summary of the results of the readability indexes by Flesch Reading Ease Score and Gunning Fog.

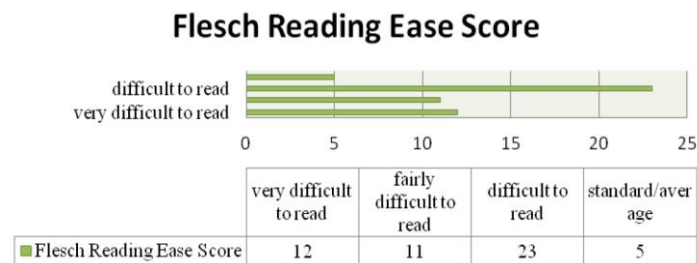


Fig. 1 The Flesch Reading Ease Score output of the texts



Fig.2 The Gunning Fog output of the texts

As Figures 1 and 2 and Table 3 represent, the majority of the texts were difficult to read. Moreover, the data analysis revealed that based on Flesch-Kincaid Grade Level, 54% of the texts were appropriate for junior and high school levels and 31% and 13% of the texts were appropriate for college and college graduate and above students, respectively. In addition, according to Coleman-Liau Index 60%, 25%, and 13% of the texts were appropriate for junior and high school, college, and graduate college students, respectively. The results of the SMOG Index, also, revealed that 28% of the texts were suitable for college students and 6% of the texts were appropriate for graduate college students.

8. DISCUSSION

The findings of the study revealed that, surprisingly enough, the majority of the texts used in the ESP textbooks published by SAMT were not appropriate for Iranian college students attending ESP courses. There has been a paucity of research with regard to studies centered on ESP textbooks. This study might provide valuable information for SAMT, materials developers, researchers, and teachers regarding the ESP textbooks used at Iranian universities context.

Unfortunately, as Wray and Janan (2013) rightly mentioned, the visibility and importance of readability has declined in the education literature. We are in urgent need of more studies investigating the readability issue. Readability indexes can provide the basic data on which one can make informed decisions about the syllabus design and developing materials (Gaies, 1979). However, as McConnell (1983) warned, placing blind faith in numbers would be simple but costly mistake; hence, readability indexes should be just one of the factors to be taken into account in developing materials. It is tempting to suggest that intended target audience is a key determinant of readability of a text (Tinkler & Woods, 2013). The readability indexes used in this study are the most commonly accepted and applied ones but it should be highlighted that these formulas are devoid of qualitative variables, namely vocabulary level, syntax and organizational of material, the cohesiveness of the discussion, the complexity of ideas and arguments, reinforcement through restatement and repetition, writing style, and discourse markers (McConnell, 1983). Also, as the current theories underscore, the individual differences, in particular motivation and aptitude and the amount of involvement of learners also play key roles in reading and comprehending a text. Moreover, due to the fact that the readability formulas were generally developed for assessing the readability of materials to be used by an individual learner in an unsupervised situation, certain factors which might lower the text difficulty for classroom have been ignored (Kerr, 1949).

No readability score per se is the complete and definite answer to assessing texts. The readability formulas have faced severe criticisms for being simplistic, relying on formal properties of a text, being based on only word length and frequency and sentence length, not being compatible with current linguistic analyses, and neglecting the semantic, pragmatic, psycho- and sociolinguistic aspects of language (Lenzner, 2014). Similarly, Oakland and Lane (2004) argued that the readability formulas do not take into account the structure-level features such as inference load or story structure that also exert an influence on text difficulty. In the same vein, Fulcher (1997) remarked that these formulas ignore other aspects of texts like font size and type, illustration and color,

specialist use of lexical items, conceptual and propositional complexity, textual organization, and syntax and factors relating to the target readers such as background knowledge, topic familiarity, interest in the subject, level of general education, reading speed and strategies, and also need of learners. Therefore, in conclusion, on the one hand, researchers need to improve the validity of readability measures and on the other hand, material developers should pay close attention to the readability of the materials adopted. The finding of this study is expected to motivate Iranian policymakers and materials developers to reconsider the ESP textbooks used at universities.

9. CONCLUDING REMARKS: THE IMPLICATIONS, LIMITATIONS, AND SUGGESTIONS

The ESP textbooks' readability studies might be invaluable for shedding more light on the quality and appropriateness of the materials used in ESP classrooms. The findings of this study and further research in this field would be useful for ESP materials developers. As the data analysis indicated, the texts used in the ESP textbooks in Iran are not appropriate for Iranian ESP undergraduate and graduate students.

A word of caution is in order; it should be highlighted that readability of texts is only one of the variables that exert an influence on the appropriateness of a textbook selection. Exclusive dependence on readability formulas might be a pitfall which should be avoided through triangulating by using other readability measurements.

Although the perceived obstacles and problems are still present in ESP courses in Iran, we hope this study might provide a renewed impetus for further research investigating critically the main problems of ESP courses in Iran, in particular the time, learners' proficiency, ESP instructors, and textbooks. Future researcher can analyze more texts and textbooks using different objective and subjective measures. Also, this study should encourage researchers to assess the readability of the ESP textbooks in other countries too.

Valid criticisms have also been leveled against readability measures; in a seminal paper, Armbruster, Osborn, and Davison (1985, p. 18) warned that "these formulas may not be very useful in selecting textbooks and that, in fact, they may adversely affect the quality of textbook writing." However, one cannot ignore readability index as an indicator of text difficulty. Benjamin (2012) claims that developments in assessing text difficulty in the past two decades lend support to the validity of some of the more traditional readability indexes. An important further step can be examining other ways for assessing the readability of texts; for instance, Crossley, Greenfield, and McNamara (2008) suggested Coh-Metrix as a cognitively-based readability index to assess text cohesion and text difficulty on different measures of language and discourse which connects the latest developments in computational linguistics and discourse processing (Crossley & McNamara, 2008) which results in multilevel analyses of textual difficulty (Elfenbein, 2011; Graesser, McNamara, Kulikowich, 2011). Therefore, another avenue for further research is to analyze the texts of ESP textbooks by Coh-Metrix.

As another option, Lexile Framework is designed which analyzes texts based on word frequency and sentence length. But, Krashen (2001) calls this framework as unnecessary and potentially harmful because it limits the available reading options for extensive reading. Similarly, as Harrison and Bakker (1998) argue, lexical density of a text and cloze procedure might be a more reliable and valid indicator of the readability of a text

than readability formulas. Begeny and Greene (2014) encouraged researchers and practitioners to go beyond readability formula in assessing text difficulty. Chavkin (1997) concluded that although teachers should be aware of readability as a criterion in textbook selection, but readability formulas are not an end in themselves. In a recent paper, Hertley (2016) recommends that we should have different readability measures for different tasks. He suggests that we could have discipline related, ability related, age related measures as well as a combination of these measures which would result in much more valid readability measure. Similarly, Chen (2016) suggested taking into account vocabulary and structure complexity in measuring text difficulty. Also, the teachers' and learners' voicing regarding the ESP syllabus and textbooks are missing too which need research.

Concisely, from the practical perspective, this study provides implications for teachers who are responsible for the textbooks they choose and use, for material developers, and for textbook publishers. Contextually speaking, most importantly, the findings of this study might provide impetus for SAMT authorities to reconsider and revise the ESP textbooks.

REFERENCES

- Atai, M. R., & Mazlum, F. (2013). English language teaching curriculum in Iran: Planning and practice. *The Curriculum Journal*, 24(3), 389–411.
- Automated readability index. (2016). Retrieved from https://en.wikipedia.org/wiki/Automated_readability_index
- Armbruster, B. B., & Anderson, T. H. (1988). On selecting considerate content textbooks. *Remedial and Special Education*, 96(1), 47–52.
- Armbruster, B. B., Osborn, J. H., & Davison, A. L. (1985). Readability formulas may be dangerous to your textbooks. *Educational Leadership*, 42, 18-20.
- Arnó Macià, E. (2012). The role of technology in teaching languages for specific purposes courses. *The Modern Language Journal*, 96(s1), 89–104.
- Bailin, A., & Greeneafstein, A. (2001). The linguistic assumptions underlying readability formulae: A critique. *Language and Communication*, 21, 285–301.
- Basturkmen, H. (2003). Specificity and ESP course design. *RELC Journal*, 34(1), 48–63.
- Begeny, J. C., & Greene, D. J. (2014). Can readability formulas be used to successfully gauge difficulty of reading materials? *Psychology in the Schools*, 51(2), 198–215.
- Benjamin, R. G. (2012). Reconstructing readability: Recent developments and recommendations in the analysis of text difficulty. *Educational Psychology Review*, 24(1), 63–88.
- Belcher, D. D. (2004). Trends in teaching English for specific purposes. *Annual Review of Applied Linguistics*, 24, 165–186.
- Belcher, D. D. (2006). English for specific purposes: Teaching to perceived needs and imagined futures in worlds of work, study, and everyday life. *TESOL Quarterly*, 40(1), 133–156.
- Brunfaut, T. (2014). English for specific purposes: Current and future issues. *Language Assessment Quarterly*, 11(2), 216–225.
- Chavkin, L. (1997). Readability and reading ease revisited: State-adopted science textbooks. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 70(3), 151–154.

- Chen, A. C. (2016). A critical evaluation of text difficulty development in ELT textbook series: A corpus-based approach using neighbor clustering. *System*, 58, 64–81.
- Cline, T. A. (1972). Readability of community college textbooks and the reading ability of the students who use them. *Journal of Reading Behavior*, 5(2), 110–118.
- Coleman–Liau index. (2016). Retrieved from https://en.wikipedia.org/wiki/coleman%e2%80%93liu_index
- Crossley, S. A., & McNamara, D. S. (2008). Assessing L2 reading texts as the intermediate level: An approximate replication of Crossley, Louwerse, McCarthy & McNamara (2007). *Language Teaching*, 41(3), 409–429.
- Crossley, S. A., Greenfield, J., & McNamara, D. S. (2008). Assessing text readability using cognitively based indices. *TESOL Quarterly*, 42(3), 475–493.
- Davies, A. (2001). The logic of testing languages for specific purposes. *Language Testing*, 18(2), 133–147.
- Elfenbein, A. (2011). Research in text and the uses of Coh-Metrix. *Educational Researcher*, 40(5), 246–248.
- Emery, H. J. (2014). Developments in LSP testing 30 years on? The case of Aviation English. *Language Assessment Quarterly*, 11(2), 198–215.
- Flesch–Kincaid readability tests. (2016). Retrieved from https://en.wikipedia.org/wiki/Flesch%E2%80%93Kincaid_readability_tests#Flesch.E2.80.93Kincaid_Grade_Level
- Fox, E. (2009). The role of reader characteristics in processing and learning from informational text. *Review of Educational Research*, 79(1), 197–261.
- Fulcher, G. (1997). Text difficulty and accessibility: Reading formulae and expert judgement. *System*, 25(4), 497–513.
- Gaies, S. J. (1979). Linguistic input in formal second language learning: The issues of syntactic gradation and readability in ESL materials. *TESOL Quarterly*, 13(1), 41–50.
- Graesser, A. C., McNamara, D. S., & Kulikowich, J. M. (2011). Coh-Metrix: Providing multilevel analyses of text characteristics. *Educational Researcher*, 40(5), 223–234.
- Grootens-Wiegers, P., De Vries, M. C., Vossen, T. E., & Van den Broek, J. M. (2015). Readability and visuals in medical research information forms for children and adolescents. *Science Communication*, 37(1), 89–117.
- Grosse, C. U., & Voght, G. (2012). The continuing evolution of languages for specific purposes. *The Modern Language Journal*, 96(s1), 190–202.
- Gunning fog index. (2016). Retrieved from https://en.wikipedia.org/wiki/Gunning_fog_index
- Harrison, S., & Bakker, P. (1998). Two new readability predictors for the professional writer: Pilot trials. *Journal of Research in Reading*, 21(2), 121–138.
- Hartley, J. (2016). Is time up for the Flesch measure of reading ease? *Scientometrics*, 107(3), 1523–1526.
- Hayati, A. M. (2008). Teaching English for special purposes in Iran: Problems and suggestions. *Arts and Humanities in Higher Education*, 7(2), 149–164.
- Heydari, P., & Riazi, A. M. (2012). Readability of texts: Human evaluation versus computer index. *Mediterranean Journal of Social Sciences*, 3(1), 177–190.
- Johns, A. M., & Dudley-Evans, T. (1991). English for specific purposes: International in scope, specific in purpose. *TESOL Quarterly*, 25(2), 297–314.
- Kalyuga, S. (2009). *Managing cognitive load in adaptive multimedia learning*. New York: Information Science Reference.

- Kasule, D. (2011). Textbook readability and ESL learners. *Reading and Writing*, 2(1), 63–76.
- Kendeou, P., Muis, K. R., & Fulton, S. (2011). Reader and text factors in reading comprehension processes. *Journal of Research in Reading*, 34(4), 365–383.
- Kerr, M. (1949). Use of readability formulas in selecting textbooks. *The Elementary School Journal*, 19(7), 411–414.
- Krashen, S. (2001). The Lexile framework: Unnecessary and potentially harmful. *CSLA Journal*, 24(2), 25–26.
- Lenzner, T. (2014). Are readability formulas valid tools for assessing survey question difficulty? *Sociological Methods & Research*, 43(4), 677–698.
- Leroy, G., Kauchak, D., & Mouradi, O. (2013). A user-study measuring the effects of lexical simplification and coherence enhancement on perceived and actual text difficulty. *International Journal of Medical Informatics*, 82, 717–730.
- Linsear write. (2016). Retrieved from https://en.wikipedia.org/wiki/Linsear_Write
- Mayer, R. E., & Moreno, R. (2010). Techniques that reduce extraneous cognitive load and manage intrinsic cognitive load during multimedia learning. In J. L. Plass, R. Moreno, & R. Brunken (Eds.), *Cognitive load theory* (pp. 131–152). New York: Cambridge University Press.
- McConnell, C. (1983). Readability: Blind faith in numbers? *The Journal of Economic Education*, 14(1), 65–71.
- McDonough, J. (2010). English for specific purposes: A survey review of current materials. *ELT Journal*, 64(4), 462–477.
- Moreno, R. & Park, B. (2010). Cognitive load theory: Historical development and relation to other theories. In J. L. Plass, R. Moreno, & R. Brunken (Eds.), *Cognitive load theory* (pp. 9–28). New York: Cambridge University Press.
- Oakland, T., & Lane, H. B. (2004). Language, reading, and readability formulas: Implications for developing and adapting tests. *International Journal of Testing*, 4(3), 239–252.
- O'Sullivan, B. (2012). Assessment issues in languages for specific purposes. *The Modern Language Journal*, 96(s1), 71–88.
- Parker, R. I., Hasbrouck, J. E., & Weaver, L. (2001). Spanish readability formulas for elementary-level texts: A validation study. *Reading & Writing Quarterly: Overcoming Learning Difficulties*, 17(4), 307–322.
- Pulido, D. (2007). The effects of topic familiarity and passage sight vocabulary on L2 lexical inferencing and retention through reading. *Applied Linguistics*, 28(1), 66–86.
- Reed, R. K., & Kershaw-Herrera, S. (2016). An examination of text complexity as characterized by readability and cohesion. *The Journal of Experimental Education*, 84(1), 75–97.
- Robison, T., Roden, T., & Szabo, S. (2015). Readability levels show that social studies textbooks are written above grade level reading. *Journal of Teacher Action Research*, 1(2), 100–112.
- Sadeghi, K., & Richards, J. C. (2016). The idea of English in Iran: An example from Urmia. *Journal of Multilingual and Multicultural Development*, 37(4), 419–434.
- Sawyer, M. H. (1991). A review of research in revising instructional text. *Journal of Reading Behavior*, 23(3), 307–333.
- Schneider, D. E. (2011). Assessing the readability of college textbooks in public speaking: Attending to entry level instruction. *Communication Teacher*, 25(4), 246–255.

- Skehan, P. (1984). Issues in the testing of English for specific purposes. *Language Testing*, 1(2), 202–220.
- Skulstad, A. S. (1999). Genre awareness in ESP teaching: Issues and implications. *Journal of Applied Linguistics*, 9(2), 285–298.
- Smog. (2016). Retrieved from <https://en.wikipedia.org/wiki/SMOG>
- Soodmand Afshar, H., & Movassagh, H. (2016). EAP education in Iran: Where does the problem lie? Where are we heading? *Journal of English for Academic Purposes*, 22, 132–151.
- Spinks, N., & Wells, B. (1993). Readability: A textbook selection criteria. *Journal of Education for Business*, 69(2), 83–87.
- Sung, Y. T., Chen, J. L., Cha, J. H., Tseng, H. C., Chang, T. H., & Chang, K. E. (2015). Constructing and validating readability models: The method of integrating multilevel linguistic features with machine learning. *Behavior Research Methods*, 47(2), 340–354.
- Sung, Y. T., Dyson, S. B., Chen, Y. C., Lin, W. C., & Chang, K. E. (2015). Leveling L2 texts through readability: Combining multilevel linguistic features with the CEFR. *The Modern Language Journal*, 99(2), 371–391.
- Swales, J. M. (2000). Language for specific purposes. *Annual Review of Applied Linguistics*, 20, 59–76.
- Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, 4, 295–312.
- Sweller, J. (2005). The redundancy principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 159–168). New York: Cambridge University Press.
- Tamor, L. (1981). Subjective text difficulty: An alternative approach to defining the difficulty level of written text. *Journal of Reading Behavior*, 13(2), 165–172.
- Tinkler, S., & Woods, J. (2013). The readability of principles of macroeconomics textbooks. *The Journal of Economic Education*, 44(2), 178–191.
- Wall, S. (1969). Readability- a neglected criterion in secondary textbook selection. *Journal of the Reading Specialist*, 9(1), 12–22.
- White, S. (2012). Mining the text: 34 text features that can ease or obstruct text comprehension and use. *Literacy Research and Instruction*, 51(2), 143–164.
- Williamson, G. L. (2008). A text readability continuum for postsecondary readiness. *Journal of Advanced Academics*, 19(4), 602–632.
- Wray, D., & Janan, D. (2013). Readability revisited? The implications of text complexity. *The Curriculum Journal*, 24(4), 553–562.
- Zamanian, M., & Heydari, P. (2012). Readability of texts: State of the art. *Theory and Practice in Language Studies*, 2(1), 43–53.