ISSN 1799-2591
Theory and Practice in Language Studies, Vol. 13, No. 1, pp. 251-256, January 2023
DOI: https://doi.org/10.17507/tpls.1301.29

# Readability Assessment of Advanced English Textbooks: A Corpus-Linguistic Study

Tunan Hu
College of Foreign Languages, Zhejiang University of Finance and Economics, Hangzhou, China

***Abstract*—As a quantitative index of the ease of reading text, readability has been used as an effective manner to assess the difficulty of reading textbooks. Based on a corpus consisting of 40 texts from two sets of advanced English textbooks published by two top-tier publishing houses in China in the same year, this paper employs three readability formulas (i.e. FRE, FKGL and LR) to examine their readability trends and differences in readability. The results show that: 1) the readability of both book sets is low, 2) no significant differences are found in three readability indices, 3) some of LR subindices, such as deep cohesion and connectivity, show an opposite tendency to the overall readability. It is claimed that: 1) both book sets are fairly difficult to read; 2) they are interchangeable in the teaching process without the absolute difference in authoritativeness; 3) the increasing trend of overall readability is moderated by controlling some subindices deliberately to keep the balance between text-reading difficulty and practical demands. This study not only benefits scholars but also teachers to evaluate and improve English textbooks.**

***Index Terms*—readability assessment, advanced English textbooks, corpus**

## I. Introduction

Advanced English is one of the core courses for Chinese undergraduate English-major students. According to the *English Teaching Syllabus for English Majors* issued by the Ministry of Education, this course is designed to train students' comprehensive English skills, especially reading comprehension, grammar and rhetoric, and writing ability, through reading and analyzing a wide range of materials. Thus, as the main source of reading materials, textbooks used in this course play an important role in the teaching process. To some degree, the quality of textbooks, especially the hierarchy of readability, potentially influences the success or failure of teaching.

Readability refers to the ease of reading or understanding certain reading or written materials (Dale & Chall, 1948; Klare, 1963; McLaughlin, 1969; Richards et al., 2006). Assessing the extent to which vocabulary and sentence structure relate to the understanding of materials, readability is helpful for teachers to select proper reading materials for learners (Sheehan et al., 2014). So far, readability formula based on particular indices, especially word length and sentence length, is the popular way to evaluate the readability of texts. Since the first readability formula was devised by Lively and Pressey (1923), scholars have been working to renew formulas to evaluate readability more accurately. More and more readability formulas have been put forward. Some of them, such as Flesch Reading Ease Readability Formula (Flesch, 1948), Automated Readability (Smith & Senter, 1967), Smog (Mc Laughlin, 1969) and Flesch Kincaid Formula (Kincaid et al., 1975), has been more widely adopted for decades. In these formulas, word length, sentence length, and percentage of hard words serve as crucial subindices of readability. Although the above indices are directly related to text difficulty, they only reflect part of the superficial features of text, that is, part of text difficulty. Recent evidence suggests that the comprehension process involves multidimensional levels of understanding (Graesser & McNamara, 2011). To overcome the limitation of traditional readability formulas, Coh-Metrix L2 Readability Formula comes into the picture, which can describe a text in accordance with its syntax, vocabulary, morphology, semantics, etc. (Crossley et al., 2008). Data from several studies show that the accuracy of Coh-Metrix L2 Readability Formula is higher than that of the previous ones (Crossley et al., 2011).

In recent years, with the continuous promotion of English major reform in colleges and universities, new versions of English textbooks emerge in endlessly. As a result, how to select and evaluate textbook becomes a major area of interest within the field of teaching practice and research. However, compared with qualitative studies in this field, far too little attention has been paid to quantitative analysis (Zhao & Zheng, 2006; Yang & Chen, 2013). Among the few quantitative studies, it is the major focus to assess readability of college English textbooks for non-English-major students with traditional unidimensional formulas (Gu & Guan, 2003; Deng, 2013). Since readability of English major textbooks has not been explored fully yet, the present study is to investigate the readability trend in this unexplored field on the basis of two sets of advanced English textbooks. More specifically, we aim to solve the following two research questions: (1) Are there any significant differences in three readability indices between the identical numbered volumes within both book sets as well as between the different volumes within the same book set? (2) Are there any significant differences in eight subindices of Coh-Metrix L2 readability between the identical numbered volumes within both book sets as well as between the different volumes within the same book set?

© 2023 ACADEMY PUBLICATION

## II. Methodology

### A. Data Collection

The reading materials were taken from two sets of advanced English textbooks (hereafter referred to as TA and TB respectively). Both of them were two-volume books published in 2013 by two top-tier publishing houses in China individually. Each volume had ten units, each containing two texts. There existed a hierarchical relation between vol. 1 and vol. 2: the former was used in the first semester of the junior year, and the latter in the second one. It implied, to some degree, that vol. 1 could theoretically be easier than vol. 2. According to the guide for faculty, we excluded texts whose types were self-learning materials after class. In other words, only the first text in each unit was included for the readability assessment. This resulted in a corpus of 40 texts (about 56,229 words in total). The number of words per volume is listed in Table 1 below.

Table 1
Descriptive Statistics Of The Corpus

| Vol. | TA-vol. 1 | TA-vol. 2 | TB-vol. 1 | TB-vol. 2 | Total |
|---|---|---|---|---|---|
| Number | 12,980 | 17,312 | 11,371 | 14,566 | 56,229 |

### B. Research Tool

In the experiment, we employed the automated tool Coh-Metrix version 3.0, developed by McNamara team at the University of Memphis (available at http://141.225.61.35/cohmetrix2017).

For the first research question, Coh-Metrix automatically calculated the readability scores of both book sets with three formulas: Flesch Reading Ease Readability (hereafter, FRE), Flesch-Kincaid Grade Level (hereafter, FKGL) and Coh-Metrix L2 Readability (hereafter, LR). The result of FRE ranges from 0 to 100, with a higher score implying easier reading. In general, the output of FKGL is a number from 1 to 18+. The lower the number, the easier it is to read to text. Both of them are based on the length of sentences and words within the text, and have their own corresponding difficult level list. In contrast, cohesive devices between sentences are considered in LR formula which does not have the corresponding difficult level list. Similar to FRE, the higher the LR score, the higher the readability, and the easier for readers to understand the text.

For the second research question, it computed scores on eight subindices of LR respectively. To be exact, they are narrativity, word concreteness, syntactic simplicity, referential cohesion, deep cohesion, verb cohesion, connectivity and temporality (McNamara et al., 2014).

### C. Research Process

After the collection of the reading materials and the preparation of readability assessment, we began our analysis in three steps.

Step 1, three readability scores of the corpus were calculated to present the overall readability trend of TA and TB.

Step 2, the differences in three readability indices between the identical numbered volumes within both book sets as well as between the different volumes within the same book set were calculated via Independent-Sample T Test.

Step 3, the differences in eight subindices of LR between the identical numbered volumes within both book sets as well as between the different volumes within the same book set were calculated via Independent-Sample T Test.

## III. Data Analysis and Discussion

### A. Overall Readability Trend of TA and TB

In this section, we focus on presenting the overall readability trend of TA and TB from two perspectives: the unidimensional one (FRE, FKGL) and the multidimensional one (LR). We calculated FRE, FKGL and LR scores of each text in our corpus as well as the average scores of each volume within both book sets. Table 2 below displays the detailed descriptive statistics of the above scores.

As mentioned above, FRE and FKGL are both based on the length of sentences and words within the text. Thus, from Table 2, it is clear that FRE and FKGL scores echo each other. More specifically, for the same text or volume, the higher the FRE score, the lower the FKGL one, and vice versa. Measured by FRE, the average scores of each volume within TA are 58.26 and 53.76, within TB are 59.90 and 51.72, respectively. According to Flesch (1948), all these scores fall into the fairly difficult level (value ranging from 50 to 60), corresponding to the reading level of $10^{th}$-$12^{th}$ grade students in the US. Similarly, calculated by FKGL, the mean scores of each volume within TA are 10.68 and 12.26, within TB are 10.30 and 12.20, respectively. In accordance with Kincaid et al. (1975), FKGL scores ranging from 10 to 12 fall into the fairly difficult level, which is consistent with the FRE difficulty level above. Assessed by LR, the average scores of each volume are 15.84, 13.77, 15.83 and 12.28, respectively.

Taken together, these results demonstrate that both book sets are fairly difficult to read, aligning with the curriculum orientation of advanced English. In each volume, text difficulty is not constant or not keeping increasing from beginning to end, but complementary. The less readable texts and the more readable ones are essentially equal in number. Compared with the constant trend or that of rectilinear rise, the complementary distribution of readability is

© 2023 ACADEMY PUBLICATION

more in line with the pace of teaching practice which should be adjusted in terms of the participants' learning needs rather than follows a specific model throughout.

Furthermore, in both book sets, the reading difficulty of vol. 2 is ranked above that of vol. 1. Ideally, it is expected that students' reading comprehension ability should increase linearly with school years. Correspondingly, the reading difficulty of the textbook should change accordingly. As mentioned earlier, according to the guide for faculty, vol. 1 and vol. 2 are used individually in the first and second semesters of the junior year. Thus, in a sense, the overall readability trend illustrates that both book sets have a scientific readability hierarchy and can achieve the course objective that gradually enhances reading comprehension ability.

TABLE 2
FRE, FKGL AND LR SCORES OF THE CORPUS

| Readability | Vol. | Unit | | | | | | | | | | Average |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | |
| FRE | TA-vol. 1 | 52.49 | 67.84 | 46.21 | 36.29 | 56.15 | 57.40 | 60.97 | 55.78 | 91.55 | 57.91 | 58.26 |
| | TA-vol. 2 | 49.73 | 17.60 | 69.14 | 57.75 | 66.40 | 56.67 | 63.16 | 52.36 | 44.06 | 60.75 | 53.76 |
| | TB-vol. 1 | 58.18 | 65.37 | 68.34 | 39.14 | 59.69 | 69.16 | 56.08 | 67.93 | 61.62 | 53.51 | 59.90 |
| | TB-vol. 2 | 56.02 | 18.29 | 48.82 | 58.77 | 51.24 | 61.33 | 22.57 | 43.05 | 88.04 | 69.03 | 51.72 |
| FKGL | TA-vol. 1 | 12.67 | 7.48 | 12.56 | 19.44 | 10.75 | 11.62 | 9.96 | 9.84 | 2.87 | 9.66 | 10.68 |
| | TA-vol. 2 | 13.86 | 24.16 | 7.04 | 10.39 | 8.64 | 13.35 | 10.29 | 12.15 | 13.20 | 9.50 | 12.26 |
| | TB-vol. 1 | 10.90 | 8.84 | 8.80 | 16.15 | 10.70 | 9.44 | 9.93 | 7.78 | 8.38 | 12.08 | 10.30 |
| | TB-vol. 2 | 12.17 | 17.91 | 12.26 | 9.91 | 11.29 | 8.68 | 25.43 | 12.82 | 3.43 | 8.10 | 12.20 |
| LR | TA-vol. 1 | 14.14 | 16.43 | 11.65 | 9.77 | 15.08 | 16.46 | 23.42 | 7.54 | 24.97 | 18.93 | 15.84 |
| | TA-vol. 2 | 17.15 | 8.36 | 9.39 | 13.32 | 10.93 | 19.13 | 16.83 | 12.09 | 11.68 | 18.77 | 13.77 |
| | TB-vol. 1 | 19.24 | 11.76 | 15.65 | 10.00 | 14.84 | 19.26 | 17.16 | 23.71 | 13.58 | 13.15 | 15.83 |
| | TB-vol. 2 | 13.96 | 8.23 | 9.71 | 10.34 | 7.83 | 17.32 | 11.85 | 10.17 | 17.90 | 15.48 | 12.28 |

### *B. Differences in Three Readability Indices*

The first research question is answered in this section. Intuitively, according to the previous analysis, larger or smaller differences exist in three readability indices between the identical numbered volumes within both book sets as well as between the different volumes within the same book set. It is necessary for us to examine whether the differences are significant statistically. The statistical approach used in this study is Independent-Sample T Test. We calculated the significance of the differences in three readability indices between the identical numbered volumes within both book sets as well as between the different volumes within the same book set separately.

Table 3 reports the results of Independent-Sample T Test of differences in three readability indices between the identical numbered volumes within both book sets. In terms of three readability indices, no statistically significant differences are detected between the scores of TA-vol. 1 and that of TB-vol. 1 (t = -0.304, 0.251, 0.002, $p$ = 0.764, 0.805, 0.998) as well as between the scores of TA-vol. 2 and that of TB-vol. 2 (t = 0.255, 0.024, 0.875, $p$ = 0.802, 0.981, 0.393). This shows that both book sets can be used interchangeably in teaching practice according to pedagogical needs. In other words, the results from the unidimensional formulas and the multidimensional one both indicate that TA and TB do not differ from each other in terms of authoritativeness.

Table 4 shows the results of Independent-Sample T Test of differences in readability between the different volumes in the same sets. Similarly, in the three types of readability, significant differences are absent between the scores of TA-vol. 1 and that of TA-vol. 2 (t = 0.687, -0.795, 0.964, $p$ = 0.501, 0.442, 0.348) and the scores of TB-vol. 1 and that of TB-vol. 2 (t = 1.148, -0.933, 2.044, $p$ = 0.266, 0.363, 0.56). As mentioned previously, to achieve the course objective improving reading comprehension gradually, the texts in vol. 2 are deliberately designed to be more difficult to read than that in vol. 1. And intuitionally, the decreasing trends of three types of readability seem to imply that both sets of advanced English textbooks have scientific readability hierarchy. However, the statistics reported in Table 4 showed the opposite result that there were no significant differences in readability between the different volumes in the same set of textbooks. This does not mean that the reading difficulty of the second volumes of both textbooks is far from the expected goal. Our previous analysis revealed that both TA and TB are fairly difficult to read. The reading materials on this difficulty level which corresponds to the reading level of $10^{th}$-$12^{th}$ grade students in the US meet the need of advanced English course. It is worth noting that the reading difficulty is not "the higher the better". As a result, the quest for high reading difficulty should not be pursued blindly, without an understanding of its relevance to the instructional objective and the

© 2023 ACADEMY PUBLICATION

instructional needs.

TABLE 3
THE RESULTS OF INDEPENDENT-SAMPLE T TEST OF DIFFERENCES IN READABILITY BETWEEN THE IDENTICAL NUMBERED VOLUMES WITHIN BOTH BOOK SETS

| Readability | Vol. 1 | | | | | Vol. 2 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | TX | M | SD | t | *p* | TX | M | SD | t | *p* |
| FRE | TA | 58.25 | 14.45 | -0.304 | 0.764 | TA | 53.76 | 14.83 | 0.255 | 0.802 |
| | TB | 59.90 | 9.09 | | | TB | 51.72 | 20.63 | | |
| FKGL | TA | 10.69 | 4.20 | 0.251 | 0.805 | TA | 12.26 | 4.73 | 0.024 | 0.981 |
| | TB | 10.30 | 2.43 | | | TB | 12.20 | 5.97 | | |
| LR | TA | 15.84 | 5.56 | 0.002 | 0.998 | TA | 13.77 | 3.92 | 0.875 | 0.393 |
| | TB | 15.83 | 4.10 | | | TB | 12.28 | 3.67 | | |

TABLE 4
THE RESULTS OF INDEPENDENT-SAMPLE T TEST OF DIFFERENCES IN READABILITY BETWEEN THE DIFFERENT VOLUMES WITHIN THE SAME BOOK SET

| Readability | TA | | | | | TB | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Vol. | M | SD | t | *p* | Vol. | M | SD | t | *p* |
| FRE | 1 | 58.25 | 14.45 | 0.687 | 0.501 | 1 | 59.90 | 9.09 | 1.148 | 0.266 |
| | 2 | 53.76 | 14.83 | | | 2 | 51.72 | 20.63 | | |
| FKGL | 1 | 10.69 | 4.20 | -0.795 | 0.442 | 1 | 10.30 | 2.43 | -0.933 | 0.363 |
| | 2 | 12.26 | 4.73 | | | 2 | 12.20 | 5.97 | | |
| LR | 1 | 15.84 | 5.56 | 0.964 | 0.348 | 1 | 15.83 | 4.10 | 2.044 | 0.56 |
| | 2 | 13.77 | 3.92 | | | 2 | 12.28 | 3.67 | | |

### *C. Differences in Eight Subindices of LR*

In this section, we focus on answering the second research question. As shown earlier, we found no significant differences in three readability indices between the identical numbered volumes within both book sets as well as between the different volumes within the same book set. Distinguished from the unidimensional formulas based on the length of sentences and words, such as FRE and FKGL, eight additional subindices are considered in LR which enables a more in-depth and comprehensive assessment of readability. To investigate further the similarities and differences between TA and TB, Independent-Sample T Test was used to compare the eight subindices of LR.

Table 5 presents the results of Independent-Sample T Test of differences in the eight subindices between the identical numbered volumes within both book sets. As specified in the table, no statistically significant differences are found in the eight subindices of LR. Here the results confirm our previous analysis that TA and TB can be used interchangeably in the teaching process.

Table 6 displays the results of Independent-Sample T Test of differences in the eight subindices between the different volumes within the same book set. As shown in it, seven of the eight subindices of LR imply that there are no statistically significant differences, and the remaining one (verb cohesion) indicates a significant downward trend in TB ($t = 2.340$, $p = 0.031$). That is, in terms of verb cohesion which is considered as a mark of texts of lower grade levels (McNamara et al., 2012), TB-vol. 2 becomes more difficult to read than TB-vol. 1. In the same index, the downward trend is insignificant in TA ($M = 0.28, 0.14$, $t = 0.500$, $p = 0.623$). However, the difference in the single index does not allow us to reach a firm conclusion about the pros and cons of the two book sets. Although vol. 2 should be more difficult to read than vol. 1 according to the original intention of the textbook design, it does not mean that reading difficulty is allowed to increase linearly without limit. As we have stated before, the blind pursuit of high reading difficulty is not desirable. The readability of textbooks should keep balance with students' reading comprehension ability and teachers' practical demands in the teaching process.

Moreover, there are some commonalities between TA and TB as seen in Table 6. In terms of syntactic simplicity, as the volume number rises, the scores somewhat decrease in both book sets. According to Biber (1988), the more complex the sentence structure is, the more words and concepts the reader needs to deal with. It follows from the above that the sentence structure in vol. 2 is relatively more difficult than that in vol. 1. In contrast, in other subindices, such as deep cohesion and connectivity both of which imply easier reading with a higher score, both book sets show the same trends that vol. 2 seems easier to read than vol. 1. These results appear to be contradictory not only to each other but also to our previous analysis. In fact, the overall change of readability does not mean that all the subindices have to change correspondingly. Moreover, as we have mentioned above, reading difficulty should be referenced to students' reading comprehension ability and teachers' practical demands in the teaching process. When necessary, overall readability can be adjusted by controlling some subindices.

© 2023 ACADEMY PUBLICATION

TABLE 5
THE RESULTS OF INDEPENDENT-SAMPLE T TEST OF DIFFERENCE IN EIGHT SUBINDICES BETWEEN THE IDENTICAL NUMBERED VOLUMES WITH BOTH BOOK SETS

| Indices | Vol. 1 | | | | | Vol. 2 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | TX | M | SD | t | p | TX | M | SD | t | p |
| Narrativity | TA | 0.03 | 0.83 | -0.100 | 0.921 | TA | 0.25 | 0.55 | 0.780 | 0.445 |
| | TB | 0.06 | 0.32 | | | TB | 0.01 | 0.84 | | |
| Syntactic simplicity | TA | -0.35 | 0.94 | 0.060 | 0.953 | TA | -0.76 | 0.83 | -0.494 | 0.628 |
| | TB | -0.37 | 0.50 | | | TB | -0.56 | 1.03 | | |
| Word Concreteness | TA | -0.20 | 0.91 | 0.038 | 0.970 | TA | -0.45 | 0.47 | -1.089 | 0.291 |
| | TB | -0.22 | 1.13 | | | TB | -0.15 | 0.73 | | |
| Referential cohesion | TA | -0.43 | 0.61 | 1.244 | 0.229 | TA | -0.42 | 0.72 | 2.101 | 0.050 |
| | TB | -0.72 | 0.43 | | | TB | -1.06 | 0.67 | | |
| Deep cohesion | TA | 0.60 | 0.90 | 1.439 | 0.167 | TA | 0.67 | 0.51 | 1.362 | 0.190 |
| | TB | 0.13 | 0.52 | | | TB | 0.26 | 0.80 | | |
| Verb cohesion | TA | 0.28 | 0.63 | -1.176 | 0.255 | TA | 0.14 | 0.64 | 0.566 | 0.578 |
| | TB | 0.63 | 0.70 | | | TB | -0.01 | 0.50 | | |
| Connectivity | TA | -2.57 | 1.20 | 0.174 | 0.864 | TA | -2.41 | 0.62 | -0.479 | 0.638 |
| | TB | -2.64 | 0.68 | | | TB | -2.26 | 0.78 | | |
| Temporality | TA | -0.10 | 0.57 | 0.358 | 0.725 | TA | -0.44 | 0.48 | -1.065 | 0.301 |
| | TB | -0.19 | 0.50 | | | TB | -0.23 | 0.37 | | |

TABLE 6
THE RESULTS OF INDEPENDENT-SAMPLE T TEST OF DIFFERENCES IN EIGHT SUBINDICES BETWEEN THE DIFFERENT VOLUMES WITHIN THE SAME BOOK SET

| Indices | TA | | | | | TB | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Vol. | M | SD | t | p | Vol. | M | SD | t | p |
| Narrativity | 1 | 0.03 | 0.83 | -0.708 | 0.488 | 1 | 0.06 | 0.32 | 0.187 | 0.855 |
| | 2 | 0.25 | 0.55 | | | 2 | 0.01 | 0.84 | | |
| Syntactic simplicity | 1 | -0.35 | 0.94 | 1.028 | 0.317 | 1 | -0.37 | 0.50 | 0.501 | 0.623 |
| | 2 | -0.76 | 0.82 | | | 2 | -0.56 | 1.03 | | |
| Word Concreteness | 1 | -0.20 | 0.91 | 0.787 | 0.445 | 1 | -0.22 | 1.13 | -0.147 | 0.884 |
| | 2 | -0.45 | 0.47 | | | 2 | -0.15 | 0.73 | | |
| Referential cohesion | 1 | -0.43 | 0.61 | -0.043 | 0.966 | 1 | -0.72 | 0.43 | 1.373 | 0.186 |
| | 2 | -0.42 | 0.72 | | | 2 | -1.06 | 0.67 | | |
| Deep cohesion | 1 | 0.60 | 0.90 | -0.206 | 0.839 | 1 | 0.13 | 0.52 | -0.429 | 0.673 |
| | 2 | 0.67 | 0.51 | | | 2 | 0.26 | 0.80 | | |
| Verb cohesion | 1 | 0.28 | 0.63 | 0.500 | 0.623 | 1 | 0.63 | 0.70 | 2.340 | 0.031 |
| | 2 | 0.14 | 0.64 | | | 2 | -0.01 | 0.50 | | |
| Connectivity | 1 | -2.57 | 1.20 | -0.366 | 0.718 | 1 | -2.64 | 0.67 | -1.167 | 0.259 |
| | 2 | -2.41 | 0.62 | | | 2 | -2.26 | 0.78 | | |
| Temporality | 1 | -0.10 | 0.57 | 1.441 | 0.167 | 1 | -0.19 | 0.50 | 0.253 | 0.803 |
| | 2 | -0.44 | 0.48 | | | 2 | -0.23 | 0.37 | | |

## IV. CONCLUSION

This research provides a readability assessment of two sets of advanced English textbooks. Different from the previous studies measuring readability through only the unidimensional formulas, the present study examined the readability of the textbooks from two perspectives: the unidimensional one (i.e. FRE and FKGL) and the multidimensional one (i.e. LR). The results first confirm the fairly difficult nature of both book sets which aligns with the curriculum orientation of advanced English. With two research questions, we also examined whether there are significant differences in three readability indices as well as in eight subindices of LR between the identical numbered volumes within two book sets and also between the different volumes within the same book set. In terms of the first research question, the results reported no significant differences in three readability indices, indicating that both book sets could be used interchangeably. For the second one, our findings show that the increasing trend of overall readability could be moderated by controlling some subindices deliberately to achieve the balance between text-reading difficulty and practical demands in the teaching process. In addition to the above findings, this study also tries to provide a quantitative approach to evaluate English textbooks in terms of readability.

## ACKNOWLEDGEMENTS

This work was supported in part by grants from “Zhejiang Province Philosophy and Social Science Planning Project” (22NDQN256YB), and “Humanities and Social Science Research Project of Ministry of Education of China” (22YJC740022).

## REFERENCES

[1] Biber, D. (1998). *Variation across speech and writing*. Cambridge: Cambridge University Press.

© 2023 ACADEMY PUBLICATION

[2] Crossley, S. A., D. Allen & D. S. McNamara. (2011). Text readability and intuitive simplification: a comparison of readability formulas. *Reading in a Foreign Language*, *23*(1): 84-102.
[3] Crossley, S. A., Salsbury, T., McCarthy, P. M., & McNamara, D. S. (2008), LSA as a measure of coherence in second language natural discourse. In V. Sloutsky, B. Love, & K. McRae (Eds.), *Proceedings of the 30th annual conference of the Cognitive Science Society* (pp. 1906–1911). Washington, DC: Cognitive Science Society.
[4] Dale, E., & Chall, J. S. (1948). A formula for predicting readability: Instructions. *Educational Research Bulletin*, *27*(2), 37–54.
[5] Deng, W. B. (2013). Xin, lao 21 shi ji da xue ying yu du xie jiao cheng dui bi yan jiu [A comparative study on new and old editions of intensive reading textbooks of 21st century college english]. *Chang chun da xue xue bao*, *23*(1): 118-121.
[6] Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, *32*(3), 221–233.
[7] Graesser, A. C., & McNamara, D. S. (2011). Computational analyses of multilevel discourse comprehension. *Topics in Cognitive Science*, *3*(2), 371–398.
[8] Gu, X. D. & Guan, X. X. (2003). CET yue du ce shi he da xue ying yu yue du cai liao yi du du chou yang yan jiu [A sample study on readability of CET reading test and reading materials of college English]. *Xi an wai yu xue yuan xue bao*, *11*(3): 39-42.
[9] Kincaid, J. P., Fishburne, R. P., Jr., Rogers, R. L., & Chissom, B. S. (1975). *Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel*. Retrieved October 11, 2022, from http://www.dtic.mil/dtic/tr/fulltext/u2/a006655.pdf
[10] Klare, G. R. (1963). *Measurement of readability*. Ames: Iowa State University Press.
[11] Lively, B. A., & Pressey, S. L. (1923). A method for measuring the vocabulary burden of textbooks. *Educational Administration and Supervision*, *9*(7), 389–398.
[12] McLaughlin, G. H. (1969). SMOG grading: A new readability formula. *Journal of Reading*, *12*(8), 639–646.
[13] McNamara, D. S., A. C. Graesser & M. M. Louwerse. (2012). Sources of text difficulty: across genres and grades. In J. P. Sabatini, E. Albro & T. O'Reilly. *Measuring up: Advances in how we assess reading ability*. Plymouth: Rowman & Littlefield Education: 89-116.
[14] McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Zhiqiang, C. (2014). *Automated evaluation of text and discourse with Coh-Metrix.* New York: Cambridge University Press.
[15] Richards, J. C., Platt, J., & Platt, H. (1992). *Longman dictionary of language teaching and applied linguistics.* London: Longman.
[16] Sheehan, K. M., Kostin, I., Napolitano, D., & Flor, M. (2014). The Text Evaluator tool: Helping teachers and test developers select texts for use in instruction and assessment. *The Elementary School Journal*, *115*(2), 184–209
[17] Smith, E., & Senter, R. (1967). *Automated readability index (AMRL-TR-66-22)*. Ohio: Aerospace Medical Re.
[18] Yang, G. & Chen, L. J. (2013). 2000 nian yi lai gao xiao ying yu jiao cai yan jiu de xian zhuang yu si kao [A study on use of reflections on skills under various interaction patterns during asynchronous online discussion]. *Wai yu yu wai yu jiaoxue*, 269(2): 16-19.
[19] Zhao, Y. & Zheng, S. T. (2006). Ji ge guo wai ying yu jiao cai ping gu ti xi de li lun fen xi [Theoretical Analysis on some foreign systems of English textbook evaluation]. *Wai yu jiao xue*, *27*(3): 39-45.

**Tunan Hu**, Ph. D, is currently a lecture of School of Foreign Languages, Zhejiang University of Finance & Economics in Hangzhou, P. R. China. His research interests include typology and cognitive linguistics.

© 2023 ACADEMY PUBLICATION