



Impact of Readability and CEFR Levels on EFL Materials

1. Golda Cherian

PhD Scholar, Banasthali Vidyapith University, Rajasthan & English Lecturer, Sultan Qaboos University,
Muscat

2. Dr. Sunil Kumar Jha

M.A Ph.D, Assistant Professor, Department of English & Modern European Languages, Banasthali
Vidyapith University

Abstract

This study investigates the difficulty level of the texts used in the exams in correlation with the difficulty level of texts in the course materials in the teaching and learning of English in the EFL (English as a Foreign Language) field. The difficulty level of the texts, specifically in Reading and Listening skills, was assessed using Flesch-Kincaid readability scores and CEFR (Common European Framework of Reference for Languages) levels. By examining the various texts prepared by the researcher for the Level 3 exams in one of the higher education institutions in Oman, the study aims to make sure that the appropriate materials were used in assessing the learners. The findings reveal noteworthy insights about the text complexity of the test materials and the teaching materials with respect to the readability scores and CEFR levels. The outcomes of this study will offer insights and practical implications to exam writers and content developers in creating appropriate texts for teaching and assessing EFL learners and thereby achieving learning outcomes.

Keywords: Readability, CEFR level, impact, EFL (English as a Foreign Language), difficulty level, Flesch-Kincaid readability, Flesch reading ease, text complexity, listening comprehension, reading comprehension, assessment, practical implications, exam writers, content developers, higher education, Oman

Introduction

The role of readability in teaching and learning has significance in providing students with materials on the appropriate level of challenge (Thomas, 2003). In EFL education, in both realms of material preparation and test preparation, it is crucial to measure the complexity of texts. Flesch-Kincaid readability score and the CEFR level are the two most commonly used criteria for assessing the complexity of text materials.

The Flesch-Kincaid readability tool provides a numerical score of text difficulty, which is called Flesch-Kincaid Grade, on the basis of factors like sentence length and word complexity. In contrast, the CEFR framework classifies materials into different levels from A1 (beginner) to C2 (proficient). This classification is based on qualitative descriptors of language proficiency.

Readability is significant since it influences how well a text can be comprehended by the reader. The article “What Is Flesch-Kincaid Readability?” published by WebFX, states that Flesch-Kincaid Grade will help educators determine if the difficulty level of the text material is right for the learners.

The idea of calculating Readability formula was first introduced by Rudolph Flesch in 1946. His finding is known as the Flesch Reading Ease. It was based on a ranking scale of 0-100, and the higher your score, the better. Low scores indicate that the text is complicated to understand. The formula he used was:

$$\text{Flesch Reading Ease} = 206.835 - 1.015 \times (\text{words/sentences}) - 84.6 \times (\text{syllables/words}).$$

Later in 1976, another formula was developed by J. Peter Kincaid. He was asked by the US Navy to recalculate the Flesch Reading Ease to help sailors read the Navy training manuals faster and understand them better (“Measure reading levels with readability indexes”, 2024). The result was a new formula, known as Flesch-Kincaid Grade:

$$\text{Flesch-Kincaid Grade: } (0.39 \times \text{average number of words per sentence}) + (11.8 \times \text{average number of syllables per word}) - 15$$

Flesch-Kincaid Grade tells you the American school grade you would need to be in to comprehend the material on the page. If the score is lower, the easier the text is to read and comprehend.

According to “Readability Metrics and Reading Skills” (Australian Manual of Style, 2024), the most commonly and widely used readability score is the Flesch–Kincaid readability score. It usually represents readability scores aligned with American school grades 3–12. Later, the scale was expanded to grades 13–16 for college students and also a 17th grade for specialized technical texts. The table below shows the classification of grades aligned with the text difficulty.

Flesch-Kincaid Reading Grade	Level	Flesch Reading Ease	Words/ sentence	Syllables/ word
4th	Very easy	90-100	8 or fewer	1.23 or fewer
5th	Easy	80-90	11	1.31
6th	Fairly easy	70-80	14	1.39
7th or 8th	Standard	60-70	17	1.47
Some high school	Fairly hard	50-60	21	1.55
High school or some college	Hard	30-50	25	1.67
College	Very hard	0-30	29 or more	1.92 or more

Table 1: A comparison of Flesch Reading Ease and Flesch-Kincaid Reading Grade

The Flesch Reading Ease score is between 1 and 100, and the Flesch Kincaid Grade Level reflects the US education system (“Flesch Reading Ease and the Flesch Kincaid Grade Level” by Readable). Both these scores are based on word and sentence lengths. They are both calculated with the same units, but the weightings for these units are different between the two tests, resulting in different readability scores. To explain more, a lower Flesch Reading Ease readability score represents a higher level of readability difficulty whereas a lower Flesch-Kincaid Reading Grade means that the text is easy to understand. For example, texts scoring under 50 for Flesch Reading Ease are ‘difficult’ or ‘very difficult’ while texts scoring 5 or below for Flesch-Kincaid Grade are ‘easy’ or ‘very easy’. In terms of difficulty, the two scores are inversely proportionate on the reading score.

The Common European Framework of Reference (CEFR) is a detailed description of learner’s language level by skill. It is a useful reference for both educators and learners. It was designed as a framework for all European languages. The CEFR was developed by the Council of Europe, an organization that promotes

co-operation among all European countries. The CEFR has three broad bands – A, B and C and each of those bands is again divided into two, giving a description of six main levels (Introductory Guide to the Common European Framework of Reference (CEFR) for English Language Teachers, 2013).

The table below shows the different CEFR Bands and their descriptors.

Level	Band	General description
Proficient user	C2	Highly proficient – can use English very fluently, precisely and sensitively in most contexts
	C2	Able to use English fluently and flexibly in a wide range of contexts
Independent user	B2	Can use English effectively, with some fluency, in a range of contexts
	B1	Can communicate essential points and ideas in familiar contexts
Basic user	A2	Can communicate in English within a limited range of contexts
	A1	Can communicate in basic English with help from the listener

Table 2: CEFR levels and their descriptors

The CEFR gives a common foundation for the elaboration of syllabi, assessments, course materials and so on across Europe. It comprehensively describes what language learners must learn, to use a language for communication and what knowledge and skills they need to develop for effective communication. The band descriptions also cover the levels of proficiency, which allow the teachers to measure their learners' progress at each stage of learning. (Purposes of the CEFR, 2024).

The CEFR is used for many purposes like developing courses and text materials, creating tests, and evaluating students' writing. Teachers use these bands to assess students' proficiency level and their progress. It is also used for teacher-training programs (Introductory Guide to the Common European Framework of Reference (CEFR) for English Language Teachers, 2013).

The IELTS stands for the International English Language Testing System. It is an English Language proficiency test accepted globally. IELTS continues to be the most popular English language proficiency test in the world for higher education and global migration. IELTS is developed and administered by the British Council in partnership with IDP Education and Cambridge Assessment English (What is IELTS? All You Need to Know About IELTS). The article "Judging the difficulty of reading texts for learners of English" by the Linguapress shows the classification of language levels based on Flesch Reading Ease, CEFR level and IELTS band.

The Table below shows the comparison of these different metrics.

Flesch Reading Ease	Level	CEFR Level	IELTS band
0-50	Very difficult (Higher education level)	C2 Mastery / Proficiency	8 ⁺ - 9
50 - 60	Fairly difficult (11th or 12th grade - final years of high school)	C1 Advanced	7 ⁺ - 8
60 - 70	Plain English - should be easy to understand by students from 14 to 15 years upwards	B2 Upper intermediate	6 ⁺ - 7
70 - 80	Fairly easy - accessible to students aged 13 upwards	B1 intermediate	4.5 - 6
80 - 90	Easy	A2 elementary	3 - 4
90 - 100	Very easy	A1 EFL beginners	1 - 2

Table 3: A comparison of Flesch Reading Scores, CEFR levels and IELTS Bands

The main objective of this study is to determine the extent of correlation between readability scores and CEFR levels of the texts in the course materials and that of the texts used in assessments. It also intends to provide recommendations for educators on using these metrics effectively.

Literature Review

Crossley et al. (2011) found that text materials are regularly made simpler for language learners by authors using a variety of techniques to make the texts more comprehensible. Readability measures are an important tool that authors use for evaluating text comprehensibility. The study examined the effectiveness of different readability formulas, including Flesch-Kincaid, in predicting text difficulty for EFL learners. The study compared the traditional readability formulas and the Coh-Metrix Second Language (L2) Reading Index, a readability formula based on psycholinguistic and cognitive models of reading, on a large corpus of texts simplified for language learners. The purpose of the study was to identify the formula that best classifies text level into advanced, intermediate, or beginner with relation to the formulas' capacity to evaluate text comprehensibility. The results of the study were favorable to the Coh-Metrix L2 Reading Index.

Francois and Miltsakaki (2012) compared the Flesch readability formula with the "classic", an emerging paradigm which uses sophisticated NLP-enabled features and machine learning techniques. They carried out experiments on a corpus of texts for French as a foreign language, yielded four main results and they found that the new Flesch readability formula performed better than the "classic" formula.

Claridge (2012), drew attention to the lack of consistency between publishers in measuring the readability of EFL materials. She highlighted the big difference in categorizing the level of reading material in the Extensive Reading Program materials.

A study by Zamanian & Heydari (2012) discussed the relation between readability level and CEFR levels, emphasizing the differences between Flesch-Kincaid readability analysis and CEFR grading methods. The study found that Flesch-Kincaid underrates the difficulty of texts with short sentences and that CEFR scores do not consider linguistic familiarity or explained vocabulary with context clues.

According to an article "Graded reader readability: Some overlooked aspects" (Gillis-Furutaka, 2015), readability is considered as the degree of comprehensibility of a text. The same article mentions that checking readability is important for graded readers used in Extensive Reading programs in the ESL & EFL field, and therefore, how the readability of graded readers is measured should be evaluated.

Sung et al. (2015) stated that selecting suitable texts for ESL and EFL learners is an important approach to enhancing motivation and learning. Therefore, the study shows why it was important for a team of experts to classify 1,578 CFL texts into their appropriate CEFR levels.

Another study by Uchida and Negishi (2018) developed a system to assign CEFR-J levels to English texts based on textual features such as sentence structure and vocabulary. They used four readability indexes, including Flesch Reading Ease, and demonstrated the correlation between these indexes and CEFR levels.

An article by Natova (2021) presents qualitative and quantitative tools to help educators to assess the text complexity or difficulty according to the Common European Framework of Reference for Languages. Using the qualitative and quantitative tools, selected texts were analyzed to represent them according to their difficulty at the various CEFR levels. Later, the validity of this analysis at B1 and B2 levels was tested by inviting students at these levels to translate the corresponding texts. At least 90% of the vocabulary of the text corresponding to their level was understood by the students. This confirmed the validity of the tools proposed in assessing text complexity according to CEFR levels, at least with regard to B1 and B2 levels.

Adelia and Kuswoyo (2023), found that the low reading performance of Indonesian students was caused by the Readability level of the texts. It was crucial for English Teachers to know whether the texts were suitable

for their students or not. The study used correlational research to know if there was any correlation between readability level and students' reading comprehension. The sample was the Tenth grade students of SMA Gajah Mada and SMAN 7 Bandar Lampung in academic year 2020/2021. Reading comprehension test was conducted among students and the researcher measured the readability of the texts using Flesch Reading Ease Formula. Later the researcher classified the Readability level using CEFR Level. The study showed that there was a positive correlation between the readability level and students' comprehension.

Rubesova (2023), conducted a comparative study to analyze the readability of English written texts at CEFR B2 level created for the second language L2 reading practice. The chosen texts were tested for readability employing seven readability formulas with the aim to contrast the genuine CEFR B2 level to the declared CEFR B2 level, while comparing results with the Flesch-Reading Ease scores. Findings demonstrated a discrepancy between the proclaimed and genuine levels of analysed written texts of the corpus, including the sample texts coming from the Council of Europe website, the official producer of the CEFR level chart.

Methodology

Although educators around the world use readability score and CEFR levels widely to check the text difficulty, the impact of these two measures on material preparation continues to be underexplored. This study tries to find out the correlation of the readability score and CEFR level of texts used in Listening and Reading exams with regard to the texts in the course materials, by analyzing the scores of a few texts. Specifically, the study investigates whether Flesch-Kincaid readability score aligns with CEFR levels of the texts in the course books for Level 3 in the higher education institutions in the Sultanate of Oman.

The Sample of the Study

For the test samples, the researcher used Reading and Listening Texts prepared by herself for the Exams in the University of Technology and Applied Sciences – Ibra for the Level 3 Mid Semester and Final exams in the first semester of the academic year 2017-2018. The researcher used her own exams to avoid plagiarism or copyright issues. These exams have been exposed to students as practice exams and so, the exams are not currently kept in the exam bank. Therefore, it will not affect the exam confidentiality and security. In a Level 3 exams, there are three Listening texts and two Reading texts. For the course book samples, the researcher selected equal number of listening and reading text samples from Pathways 2 Listening, Speaking and Critical Thinking and Pathways 2 Reading, Speaking and Critical Thinking by Cengage Learning. The books are rated as CEFR B1-B2. These books were the Course materials used for Level 3 students in the University of Technology and Applied Sciences – Ibra, which was previously known as Ibra College of Technology.

Tools Used

There are different tools to measure the readability, but the study mainly concentrates on the Flesch Reading Ease and Flesch-Kincaid readability score since they are the most commonly used tools by educators. Understanding the relationship between readability scores and CEFR level can help educators choose and prepare materials to teach and test that are appropriately challenging for students at different proficiency levels of English language.

This study analyzes a variety of texts to check their readability scores and CEFR levels, using online tools like 'Readability Scoring System' (<https://readabilityformulas.com/readability-scoring-system.php>) for Readability scores, and The Text Analyzer (<http://www.roadtogrammar.com/textanalysis/>) for detecting CEFR levels.

Analysis of the texts

Appropriate Difficulty Level for Level 3 Exams Based on the Course Materials: Readability Score: 80-60, CEFR: B1-B2								
L3 Mid Semester Exam S1 2017-18			L3 Final Exam S1 2017-18			L3 Pathways Books Pathways 2 CEFR B1-B2		
CEFR Level	Flesch Reading Ease	Flesch-Kinkaid Reading Grade	CEFR Level	Flesch Reading Ease	Flesch-Kinkaid Reading Grade	CEFR Level	Flesch Reading Ease	Flesch-Kinkaid Reading Grade
Listening 1 B2 Intermediate	71 Fairly Easy	7th Grade	Listening 1 C1 Upper Intermediate	69 Standard (Plain English)	8th & 9th Grade	Listening Sample 1 (U10 INJAZ & Entrepreneurs) B2 Upper Intermediate	62 Standard (Plain English)	8th & 9th Grade
Listening 2 C1 Upper Intermediate	62 Standard (Plain English)	8th & 9th Grade	Listening 2 C1 Upper Intermediate	72 Fairly Easy	7th Grade	Listening Sample 2 (U2 Tips for Saving Energy and Protecting Our Environment) B2 Upper Intermediate	68 Standard (Plain English)	8th & 9th Grade
Listening 3 C1 Upper Intermediate	64 Standard (Plain English)	8th & 9th Grade	Listening 3 B1 Elementary	69 Standard (Plain English)	8th & 9th Grade	Listening Sample 3 (U4 Australia's Water) B2 Upper Intermediate	60 Standard (Plain English)	8th & 9th Grade
Reading 1 B2 Intermediate	64 Standard (Plain English)	8th & 9th Grade	Reading 1 C1 Upper Intermediate	63 Standard (Plain English)	8th & 9th Grade	Reading Sample 1 (U2 The Power of Creativity) B2 Intermediate	71 Fairly Easy	7th Grade
Reading 2 B2 Intermediate	61 Standard (Plain English)	8th & 9th Grade	Reading 2 B2 Intermediate	76 Fairly Easy	7th Grade	Reading Sample 2 (U4 Internet Island) B2 Intermediate	69 Standard (Plain English)	8th & 9th Grade

Table 4: Analysis of Samples from L3 Exams & L3 Books on FK Reading Scores & CEFR levels

Observations

1. All tested exams of L3 and the samples from the textbooks correspond to the suggested readability score.
2. The CEFR Levels do not always correspond to the readability scores. As it is evident from the table, reading score 69 is evaluated as C1 & B1. Similarly, another example, reading score 71 is evaluated as B2 whereas it should be B1. According to the comparison table (Table 3), 60-70 is B2, 70-80 is B1.

Limitations of These Tools

1. The Flesch-Kincaid Readability tool works on the basis of sentence length and syllable count, to provide the reading ease score and grade for a text. It does not consider contextual or semantic aspects like the complexity of words used.
2. CEFR level checkers consider the complexity of vocabulary, grammar, and context. However, CEFR levels can be subjective, because the tools depend on the description of the language levels and they do not have standardized guidelines for all types of texts or different genres.

Implications for Test Developers and Content Creators

1. Test developers and content creators should combine both metrics to ensure the difficulty level of the texts. Using both tools, readability scores and CEFR levels, can deliver a more holistic view of text difficulty.
2. Discrepancies can occur between the results of CEFR level checkers and readability score checkers. Hence, it is important to create test specifications for developing tests with clear criteria for which tool is to be used as the main tool to check test difficulty. Consequently, test developers can ensure consistency with regard to one aspect.
3. Test developers and educators should be trained to understand the strengths and limitations of each metric, ensuring more accurate text selection.

Conclusion

The study reveals that readability scores and CEFR levels indicate the difficulty level and the complexity of materials. Although both readability scores and CEFR levels offer valuable insights, depending solely on one metric might not provide a complete picture of text difficulty. A combined method, supplemented by educator judgment and learner performance, can help in creating appropriate materials for teaching and assessing EFL learners. Regular training for educators and continuous review of materials are essential to ensure the effectiveness of these metrics in educational settings.

Moreover, how the complexity of materials with respect to readability scores and CEFR levels affects students' comprehension and their achievement in exams is not explored in this study. Further studies on these aspects could provide more insights for material development and test writing.

References

1. Adelia, Rizca Kuswoyo. *The Correlation Between Readability Level and Student's Reading Comprehension in Narrative Reading Texts Of Tenth Grade*. Diss. Uin Raden Intan Lampung, 2023.
2. Claridge, Gillian. "Graded readers: How the publishers make the grade." *Reading in a Foreign Language*, vol. 24, no. 1, 2012, pp. 106–119.
3. Crossley, Scott A., David B. Allen, and Danielle S. McNamara. "Text Readability and Intuitive Simplification: A Comparison of Readability Formulas." *Reading in a Foreign Language*, vol. 23, no. 1, Apr. 2011, pp. 84-101.
4. "Flesch Reading Ease and the Flesch Kincaid Grade Level." Readable, <https://readable.com/readability/flesch-reading-ease-flesch-kincaid-grade-level/>.
5. François, Thomas, and Eleni Miltsakaki. "Do NLP and machine learning improve traditional readability formulas?" *Proceedings of the First Workshop on Predicting and Improving Text Readability for Target Reader Populations*, 2012.
6. Gillis-Furutaka, Amanda. "Graded reader readability: Some overlooked aspects." *Journal of Extensive Reading*, vol. 3, no. 1, 2015, <http://jalt-publications.org/jer/>.
7. Natova, Iva. "Estimating CEFR reading comprehension text complexity." *The Language Learning Journal*, vol. 49, no. 6, 2021, pp. 699-710.

8. "Introductory Guide to the Common European Framework of Reference (CEFR) for English Language Teachers." *Cambridge University Press*, 2013, <https://www.englishprofile.org/images/pdf/GuideToCEFR.pdf>.
9. "Measure Reading Levels with Readability Indexes." *Wylie Communications*, 2021, https://www.wyliecomm.com/2021/11/measure-reading-levels-with-readability-indexes/#_ftn2.
10. "Purposes of the CEFR." *Council of Europe*, 2024, <https://www.coe.int/en/web/common-european-framework-reference-languages/uses-and-objectives>.
11. Rubesova, S. "Genuine Readability Level of Alleged English CEFR B2 Written Texts Aimed for Second Language Learning." *EDULEARN23 Proceedings*. IATED, 2023.
12. Sung, Y., Lin, W., Dyson, S. B., Chang, K. E. & Chen, Y. "Leveling L2 texts through readability: Combining multilevel linguistic features with the CEFR." *The Modern Language Journal*, vol. 99, no. 2, 2015, pp. 371-391.
13. Thomas, George. "The Role of Readability in Today's Classrooms." *Topics in Language Disorders*, vol. 23, 2003, pp. 175-189.
14. Uchida, Shingo, and Masashi Negishi. "Assigning CEFR-J levels to English texts based on textual features." *Proceedings of the 4th Asia Pacific Corpus Linguistics Conference*, 2018, pp. 463-468.
15. "What is IELTS? All You Need to Know About IELTS." *British Council*, <https://takeielts.britishcouncil.org/take-ielts/what-ielts>.
16. "What Is Flesch-Kincaid Readability?" *WebFX*, <https://www.webfx.com/tools/read-able/flesch-kincaid/>.
17. Zamanian, Mahmood, and Pourya Heydari. "Readability of texts: State of the art." *Theory and Practice in Language Studies*, vol. 2, no. 1, 2012, pp. 43-53.
18. "Readability Metrics and Reading Skills." *Australian Manual of Style*, 2024.

