

This work is licensed
under a Creative
Commons Attribution
4.0 International
License.

Utilizing an AI-CEFR checker: What is the efficacy for utilizing CEFR results for first year university students?

Simon D.C. Townsend

Fuji University

townsend@fuji-u.ac.jp

Andrew C. Tappenden

Iwate University

andrewt@iwate-u.ac.jp

We investigated a text analyzer tool that provided Common European Framework of Reference for Languages (CEFR) level results for supporting language activities in two communication classes. In one class, students self-evaluated their language in formative assessment (FA) tasks and in the end, completed a survey of their experiences, which we evaluated. Students demonstrated insights into the thought process behind their written language and their motivation to evaluate their written English. They did this with the aim of improving their own language use and developing their language to achieve CEFR-J (Japanese version) based objectives. In class two, we applied the analyzer to video presentation transcripts for determining summative assessment (SA). Findings showed future potential for using this tool for grading and designing resources for CEFR-J activities and level-specificity. We concluded that artificial intelligence (AI)-linked CEFR-J based activities and task-based learning showed potential for further AI-language research.

Keywords: CEFR-J, CEFR checker-text analyzer, self-evaluation, language use and development, assessment

Introduction

To assist language educators in utilizing AI applications in the classroom, under strict pilot study conditions (Given, 2016) this research evaluated the use of the Cathoven AI-CEFR Checker as part of language learning for first year students in two communication classes. We primarily investigated the appropriateness of this tool for our

students (Japanese University) to independently acquire and interpret feedback for language use, and to aid them in the achievement of class objectives.

We investigated this technology and measured its efficacy in language tasks derived from our CEFR-J chart for our A2 level learners. We implemented the tool for homework tasks, with classroom discussions held to support independent learning. A survey collected learner experiences, and we evaluated their responses to determine how well CEFR feedback results supported language tasks and activity objectives. We also applied the tool at the end of the semester to consider course design incorporating speaking activities, feedback, and assessment. We endeavor to contribute AI-CEFR based analysis and related concepts based on our findings within standardized teaching practices in Japan, now and in the future. The terms “CEFR Checker” and “CEFR Analyzer” and “the tool” will interchangeably be used in this study.

CEFR and CEFR-J

In 2008, The Common European Framework of Reference - Japan (CEFR-J) established categories within the existing A1/A2 categories of the CEFR (Tono, 2016) and a word list. Our institution derived a Can-Do List from the CEFR-J, and so we needed to consider the relevant expressions and vocabulary of each descriptor (Tono, 2016). Software exists to establish the CEFR level of English passages. Uchida and Negishi (2018) note, however, that tools such as Lexile Analyzer and Coh-Metrix are not clearly linked to the CEFR-J. In this case Uchida and Negishi’s CEFR-J based tool, the Vocabulary Level Analyzer (CVLA), analyzes English passages similarly to the tool in our study. However, in their 2018 paper, there is no discussion of the classroom usage of their software, so we intend to cover this. Furthermore, as Nagai and O’Dwyer (2011) write of CEFR in course design, CEFR should provide learners with autonomy, which became part of our formative feedback strategy.

Technology in English classrooms at Japanese universities

Students view technology in classrooms as normal and beneficial (Caldwell, 2018). Many researchers of technology in the classroom examine learning management systems and vocabulary applications. Hulse (2019) discusses Google classroom, implying that students quickly adopt it. Dizon (2016) found Quizlet intuitive for students. Runnels (2016) researched students’ CEFR-J based self-assessment, based on Can-Do list objectives. However, research on usage of AI technology in self-assessment and CEFR generated discourse analysis appears to be lacking. Therefore, our aim for this study is to examine how the CEFR checker AI tool in particular might contribute to student and teacher analysis of text samples.

Method

A pilot study is small-scale and designed to test the approaches that will be used in the final study (Given, 2016). We are using Givens’ attributes to A) Test equipment (AI-CEFR Checker), in connection with B) Participant type (Learners and learning

context), though we felt that it was not relevant to investigate all seven of Given's attributes (Given, 2016, p.112). We initially considered AI's potential for student self-evaluation and the first aim (Part A) was to examine student use of the tool for reflective and formative feedback in homework tasks. The second aim (Part B) involved a teacher's use of the tool for SA, implementing it as a regular part of students' language practices (formative) in the future, and for collection of scores used in students' grading (summative) (Garrison & Ehringhaus, 2007).

CEFR Checker (analyzer)

We used Cathoven, an AI tool designed for education. It determines the difficulty of texts, analyzes text characteristics, and more (Tac, 2022). We used the CEFR Checker, which analyzes up to 500 words. The results give a CEFR estimation (see Appendix A) and scores based on vocabulary, speech-parts (verbs), and grammar variety. In the setup, students created accounts and received instructions on accessing the CEFR Checker, inserting text, and deciphering the results. Students understood they would adopt the tool for homework and project work. Teachers led regular discussions with students supporting them.

Data collection and analysis

Questionnaire responses of Part A are listed below, in the Results section. Q1–6 related to access and use of the CEFR Analyzer, students selecting ranges of frequency: 0% to 25%; 25% to 50%; 50% to 75%; and 75% to 100%. Q7–10 ascertains experiences, allowing observations of the students technical and learning processes, coupled with comments. Part B discusses the efficacy of provided CEFR scores for grading (SA) and course integration. Broad observations are narrowed in search of information relevant to the research question (Creswell, 2016) in the process of inputting student transcripts into the analyzer to obtain CEFR-J scores.

Research question

What is the efficacy for utilizing an AI-CEFR Checker and results in ESL for first year university students?

Results

Study part A: Survey

Table 1. Utilization of the tool conducted by students

	Ranges of frequency			
	0% – 25%	25% – 50%	50% – 75%	75% – 100%
Q1. I understood how to access the CEFR Analyzer	0	1	6	16
Q2. I understood how to add text into the CEFR Analyzer and get feedback results	0	2	5	16
Q3. At the end of the course-activity did you understand how to read your CEFR Results?	0	3	9	11
Q4. How often did you edit your text in the editor (left column) to improve your English?	0	3	12	8
Q5. How often did you use the CEFR Analyzer?	0	5	15	3
Q6. The CEFR results helped me to improve my English and language objectives	0	2	10	11

Table 2. Feedback results including vocabulary and general level

	Vocabulary	General level	All results
Q7. Which part of the results was most helpful feedback?	4	6	13

Note. Feedback results were presented into Vocabulary and General Level including scores for vocabulary, verb forms and sentences (see Appendix 1)

Table 3. Devices preferred by students when using the tool

	PC	Smartphone	PC – smartphone	Tablet
Q8. How I used the CEFR Analyzer	13	0	10	0

Table 4. Instruction for using the tool more effectively for language use and development

	Yes	No
Q9. I would like more instruction (help) about how to use the CEFR Analyzer	17	6

Table 5. Continuation for using the tool in future English classes

	Yes	No
Q10. I would like to use the CEFR Analyzer again in the future	21	2

We recognized two main comment threads: skills-based areas for improving English and application for language use. For example, “CEFR improved my English skills” and “I was able to know what is lacking in my English skills by using the CEFR” are comments indicating students utilized feedback results for development. Additionally, “I’m glad that I was able to learn how to make PowerPoint in addition to CEFR vocabulary” and “I was able to learn how to use Zoom (for a presentation) and try to make English sentences using difficult words using CEFR” represents students connecting self-evaluation techniques to CEFR-J based language objectives.

Study part B

We also tested CEFR Checker results as a means for grading student’s speeches, although we did not apply these to students’ actual final grades. As a result, a method of providing assessment for transcripts was recorded.

1. Students record videos of their speeches.
2. Presentations are transcribed through speech-to-text, and transcripts are reviewed by the instructor while they watch the recordings.
3. Transcription is saved to a text document.
4. Copy & paste text into the text analyzer and submit.
5. Prepare the CEFR results for evaluation.
6. Download color coded text (left column) from the analysis screen.
7. Save the text to a local student file.
8. Take a screenshot of the overall scores (left column and right column) and record.

Although simple, we found this interesting as a means of objectively identifying variety in vocabulary in student speeches. However, the main issue we encountered with this approach was how speech-to-text, in this case the built-in feature of the Microsoft software Flip, struggled with the at times heavy accents of our learner’s language. We believe, however, that this approach might open the door for further research. For example, we might investigate whether the vocabulary utilized by students in pursuit of a high CEFR Checker score is truly at level for them, if they are unable to pronounce it in such a way as to be understood by mainstream speech-to-text software.

Discussion and implications

In the survey (Q1–6), most students claimed to have quickly learned and applied various techniques, suggesting that students were forming strategies to improve their language. One advantage was its use for specific topics with key words and for language identification. The comments indicated students used feedback in diverse ways, reflectively using it for individual needs. Future research might include school-provided tablets for classroom use. Part B explored the tool’s usage for SA and looked promising for further course design, including class activities and assessment methods.

To what degree individual students used their results for making improvements was unclear. We cannot quantify improvements, nor do we understand learner beliefs of self-improvement at this stage of the study. The processes of students were not discussed and evaluated in this study. AI-CEFR results were not integrated into final scores and other assessment methods at this stage. We intended the building of language to aid overall development and have now considered deriving a task-based rubric including CEFR levels as part of learner assessment. An issue raised was identifying individual student progress within the CEFR-J and overall performance with our institution's Can-Do List. However, we believe the tool might offer a broader view of vocabulary and grammar exposition than memorization-based methods.

Conclusion

We investigated whether students could implement AI into learning, in which case, our initial results indicated that independent CEFR analysis supported individuals in language and learning development. The tool allows self-assessment inside and outside the classroom. It offers implementation of tasks based on formative and summative assessment that clearly connect to the CEFR-J. In this study, we conclude that the tool's best use appeared to be in frequent formative self-assessment carried out by students. There is great potential for further application and research linking AI generated results and CEFR-J based tasks.

References

- Caldwell, M. (2018). Japanese university students' perceptions on the use of ICT and mobile-learning in an EFL setting. *CALL-EJ*, 19(2), 188–216.
- Creswell, J. W. (2016). 30 essential skills for the qualitative researcher. Sage Publications.
- Dizon, G. (2016). Quizlet in the EFL classroom: Enhancing academic vocabulary acquisition of Japanese university students. *Teaching English with Technology*, 16(2), 40–56.
- Garrison, C., & Ehringhaus, M. (2007). *Formative and summative assessments in the classroom*. Association for Middle Level Education.
- Given, L. M. (2016). *100 questions (and answers) about qualitative research*. Sage Publications.
- Hulse, R. (2019). The use and implementation of Google Classroom in a Japanese university. *The Center of the Study of English Language Teaching Journal*, 7, 71–105.
- Nagai, N., & O'Dwyer, F. (2011). The actual and potential impacts of the CEFR on language education in Japan. *Synergies Europe*, 6, 141–152.
- Runnels, J. (2016). Self-assessment accuracy: Correlations between Japanese English learners' self-assessment on the CEFR–Japan's Can Do statements and scores on the TOEIC®. *Taiwan Journal of TESOL*, 13(1), 105–137.
- Tac, E. (2022, July 28). Home. *Cathoven AI*. <https://www.cathoven.com/en/>
- Tono, Y. (2016). The CEFR-J and its impact on English language teaching in Japan. In *JACET International Convention Selected Papers*, 4, 31–52.

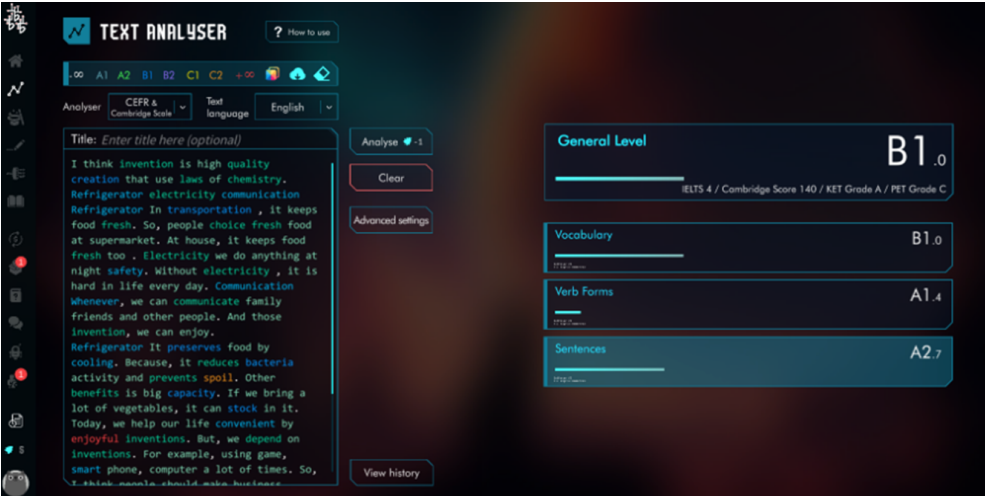
Uchida, S., & Negishi, M. (2018). Assigning CEFR-J levels to English texts based on textual features. In *Proceedings of the Asia Pacific Corpus Linguistics Conference*, 4, 463–467).



JALTCALL
Trends
Vol. 1
2024 JALTCALL
Conference Papers
Special Issue

Appendix A

CEFR feedback results – student example



Author biodata

Simon Townsend's research interests are centered on collaborative learning strategies, which encompasses technology. He is particularly interested in freeware applications due to their advanced nature and adaptability for education.

Andrew Tappenden's research interests primarily focus on the classroom experiences of his students. He aims to make language classrooms and curricula beneficial to all involved.

Swier: Utilizing an AI-CEFR checker