CEFR JOURNAL—RESEARCH AND PRACTICE

Coming Full Circle —From CEFR to CEFR-J and back

Yukio Tono, Tokyo University of Foreign Studies

https://doi.org/10.37546/JALTSIG.CEFR1-1

The CEFR-J project was launched in Japan in 2008. The CEFR-J gives a set of Can Do descriptors for 10 CEFR sub-levels (Pre-A1 to B2.2) and related Reference Level Description (RLD) work, whilst including developed profiling for vocabulary, grammar, and textual features were developed. In this article, the English resources created for the CEFR-J are applied in preparing teaching resources for other major European languages as well as Asian languages. To achieve this, a series of teaching/learning resources including the CEFR-J Wordlist and Phrase List initially developed for English were translated into 27 other languages using neural machine translation. These translated word and phrase lists were then manually corrected by a team of language experts. The automatic conversion of English to other languages was evaluated against human judgments as well as frequency analysis referencing web corpora. Three types of e-learning resources were created, taking into consideration the wordlists and the phrase lists for teaching those languages to undergraduate students: (1) a flash-card app for learning vocabulary, which allows for classification by both thematic topic and CEFR level, (2) an online syntax writing tool for the study of grammar and vocabulary, and (3) an online spoken and written production corpus collection tool.

Keywords: CEFR-J, multilingual resources, e-learning, machine translation, automatic conversion, NLP, multilingual corpora, web-based, writing tool, spoken production

1 Introduction

The Common European Framework of Reference for Languages (CEFR) was published in 2001 (Council of Europe, 2001). The CEFR is a common framework for learning, teaching and assessing a given foreign language. It features six levels (A1, A2, B1, B2, C1, and C2) on the vertical axis and skill areas (reception, interaction, production and mediation) on the horizontal axis. Commonly, these skill areas consist of Listening, Reading, Spoken Interaction, Spoken Production and Writing[1]. The framework has a third dimension, which involves other aspects of communicative competence, such as sociolinguistic, pragmatic, and strategic competences.

With the growing influence of the CEFR beyond Europe, people working in foreign language teaching and learning, notably in a number of Asian countries (Japan, Vietnam etc.), have started to explore the potential of the CEFR in their fields. The most important impact of which has been made in the area of language testing. Many foreign language proficiency tests are aligned to the respective CEFR levels and claim to be mutually comparable. As of August, 2018, the certificates of more than 30 languages are aligned to the CEFR levels according to Wikipedia[2].

In 2008, we launched a project called the CEFR-J to compile our own original framework based on the CEFR for English language teaching in Japan (Negishi, Takada, and Tono 2013; Tono 2013; Negishi

1. The self-assessment grid of the 2001 version has only one area in writing, whereas the 2018 companion volume divides writing into written interaction and written production.
2. http://en.wikipedia.org/wiki/Common_European_Framework _of_Reference_for_Languages

and Tono 2016). Some of the unique features of the CEFR-J are (1) more refined sub-levels of the CEFR (Pre-A1, A1.1-1.3, A2.1-2.2, B1.1-1.2, B2.1-2.2) with newly created and scaled descriptors, (2) the preparation of grammar and vocabulary to go with each CEFR-J level, (3) the analysis of text features to represent the CEFR-J levels, and (4) the development of tasks and tests to serve each CEFR-J descriptor (Tono 2017). The first version of the CEFR-J was released in March 2012 and is publicly available both for research/teaching and commercial purposes. The CEFR-J has been widely used as a supplement to the CEFR in Japan. The CEFR Companion Volume published in 2018 revised the framework by adding Pre-A1 and plus levels to A2, B1, and B2 respectively, which has similarities to the structure of the CEFR-J.

2 The CEFR-J x 28 project

The CEFR-J x 28 is a programme of the Super Global University (SGU) program at Tokyo University of Foreign Studies (TUFS). TUFS is a national university specialising in foreign language and culture studies, where we offer 28 different foreign languages as undergraduate majors. The number of foreign languages offered at TUFS for general education purposes exceeds 80, out of which 28 foreign languages stand as an independent major.

Despite a long history of teaching many European and Asian languages at TUFS, there was no coherent or systematic framework for teaching languages and assessing the outcomes of our program. The recent development of the CEFR and its related resources was quite inspiring to us in the sense that they offer an opportunity to systematize our teaching/learning environment by critically evaluating the current situation against a common framework. Because I have been working as a principal investigator of the CEFR-J project for English, the university thought this is a good expertise and environment to extend the research to other languages and launch the 'CEFR-J x 28' project.

This paper is an interim report on the CEFR-J x 28 project and discusses the value of constructing pedagogical resources shared across different languages, whilst examining how to best develop such resources using NLP technologies. First, a description of Reference Level Descriptions (RLDs) for English will be made (3.), and then the method of mapping the resources to multiple languages will be described (4. and 5.). Finally, as an application of the pedagogical resources, the development of three e-learning tools will be discussed and the prototype versions will be described in detail (6.).

3 CEFR-J RLD work for English

3.2 Reference Level Descriptions

The CEFR is potentially applicable to any language and does not, therefore, relate to any specific one. However, textbook authors, syllabus designers and language teachers have found its specifications to be lacking in precision, due to the language-independent nature of the framework. Consequently, Reference Level Descriptions (RLDs) have been drawn up language by language to provide reference descriptions based on the CEFR for individual languages.

The Council of Europe website on RLDs explain the details as follows: "These RLDs are made up of 'words' of a language rather than general descriptors. Reference levels identify the forms of a given language (words, grammar and so on), mastery of which corresponds to the competences defined by the CEFR. They transpose the CEFR descriptors into specific languages, level by level, from A1 to C2[3]."

According to the Council of Europe website[4], RLDs are currently available for the following languages: Croatian, Czech, English, German, French, Italian, Portuguese and Spanish. With regard to English, there are a few distinct projects related to RLDs. The English Profile (Hawkins and Filipović 2012) was an official

3. https://www.coe.int/en/web/common-european-framework-reference-languages/reference-level-descriptions (accessed August 15, 2018).
4. The same as the URL in footnote 4.

RLD piece of research carried out by a team consisting of Cambridge University, Cambridge English Assessment, Cambridge University Press, and University of Bedfordshire[5]. There are however more simplified content specifications provided by the British Council and EAQUALS in the *Core Inventory for General English* (North, Ortega and Sheehan, 2010). In addition to these academic projects, Pearson (a publishing company) developed its original scale called *Global Scale of English* (GSE)[6], which extends the CEFR by pinpointing on a scale from 10 to 90. The GSE also developed competence and performance needs to be achieved in the four skills of speaking, listening, reading and writing within a CEFR level, using a more granular approach. Furthermore GSE also provides its unique Teacher Toolkit[7], which contains 2,000 GSE learning objectives, 450 grammar objectives, and vocabulary (39,000 words and 80,000 collocations) ordered by GSE scores.

3.2 The CEFR-J RLD project

After the release of the CEFR-J version 1 in 2012, we also started to prepare RLDs for the CEFR-J in three major areas: (i) vocabulary, (ii) grammar and (iii) text properties.

3.2.1 The CEFR-J wordlist

In order to develop the wordlists for the CEFR-J, a frequency analysis of English textbooks used at primary and secondary schools in nearby Asian countries/regions (e.g. China, Korea, and Taiwan) were closely examined. The textbooks were not specifically designed based on the CEFR, but the approximate CEFR levels of the textbooks were assessed by analysing the learning objectives described in their national curriculums. In this way, we prepared Pre-A1 to B2 level sub-corpora, each of which comprised of textbook data. In the analysis of CEFR-level textbook corpora, the texts were first tagged for parts of speech (POS), using TreeTagger (Schmidt 1994) and then the frequency lists of lemmas with POS were created for each textbook published in each country/region as well as each CEFR level. Finally, the Pre-A1 words were determined by selecting only the words which appeared in all three regions' textbooks classified at the Pre-Al level. The A1-level words were then extracted in the same way, after subtracting all the Pre-A1 words from the texts in advance. In this way, vocabulary for each CEFR level was determined. Interestingly, since the vocabulary growth between Pre-A1 and A1-levels was very small (only 100 words), the two levels were merged into A1-level. Table 1 shows the breakdown of the wordlist. The 'Corpus' row indicates the initial query results of the words found across all the three regions' textbooks at a given level. The third row shows our initial target number of words. Altogether we expected to have 6,000 words from A1 to B2 levels, but after the analysis of textbook corpora, we compared our results with the English Vocabulary Profile (EVP) compiled by the English Profile team and found that while the first two levels (A1 and A2) cover a relatively homogeneous set of words, there is a larger gap in B1 and B2 level words between the two lists, so we decided to incorporate those words missing from our list, but exist in the EVP. The row called 'Final Version' shows the number of entries in the final version of the wordlist.

Table 1. ***The breakdown of the CEFR-J Wordlist***

Level	A1	A2	B1	B2	Total
Corpus	976	1,057	1,884	1,722	5,639
Our initial target	1,000	1,000	2,000	2,000	6,000
Final Version	1,068	1,358	2,359	2,785	7,570

5. The English Profile page (http://www.englishprofile. org/)
6. https://www.pearsonelt.com/about/gse.html
7. https://www.english.com/gse/teacher-toolkit/user/lo

The final version of the wordlist was then annotated with the notion categories from the *Core Inventory for General English* (North, Ortega & Sheehan 2010) and *Threshold Level* (van Ek and Trim 1990), which enables the users to extract level-appropriate vocabulary belonging to a particular thematic category. Table 2 shows a sample list of entries from the CEFR-J Wordlist.

The CEFR-J Wordlist was made publicly available in 2012. Access to the wordlist can be found on the resource page of the CEFR-J website[8]. This wordlist will serve as one of the important resources for the CEFR-J x 28 project later on.

Table 2. ***The entries of the CEFR-J Wordlist***

Entry	CEFR level	POS	Thematic domains
activity	A1	n	Leisure activities
actor	A1	n	Work and Jobs
age	A1	n	Personal information
airplane	A1	n	Ways of travelling
airport	A1	n	Travel and services vocab
animal	A1	n	
answer	A1	n	
apple	A1	n	Food and drink
apron	A1	n	Objects and rooms

3.2.2 The CEFR-J Grammar Profile

In the JSPS KAKEN project (Kiban A; No. 24242017; 2012-15), we conducted RLD research similar to previous projects such as the English Profile or the Core Inventory. There were two reasons why we had an independent RLD project. First, the CEFR-J has many sub-levels below A1 to B2, and it was desirable to specify grammar and vocabulary to go with each sub-level. For this purpose, the resources provided by the English Profile or the Core Inventory were not sufficient. Second, past reports on RLDs did not always specify the procedure of how each item of grammar or vocabulary had been assigned to a given CEFR level. Overall methods were presented, but they did not make the actual data available. Thus, we had a genuine methodological interest in how to produce RLDs in an objective, valid way. We aimed to be as transparent as possible throughout all the stages of RLD work, and made sure that the procedure would be available as a standard for those who wish to work on their own RLD research. In addition, we used corpus-based approaches similar to that of English Profile, albeit our profiling technique was very different from theirs, which would be methodologically interesting to compare.

In our project, identification of the CEFR levels was considered a type of classification task defined in the field of Natural Language Processing (NLP). Figure 1 illustrates this point. In short the classification involves supervised learning of features in the texts with the CEFR level information. First, a machine creates a certain model based on a set of feature vectors from training texts with some class information, such as CEFR levels. Then the model predicts a CEFR level when a new text is given.

8. http ://www.cefr-j.org

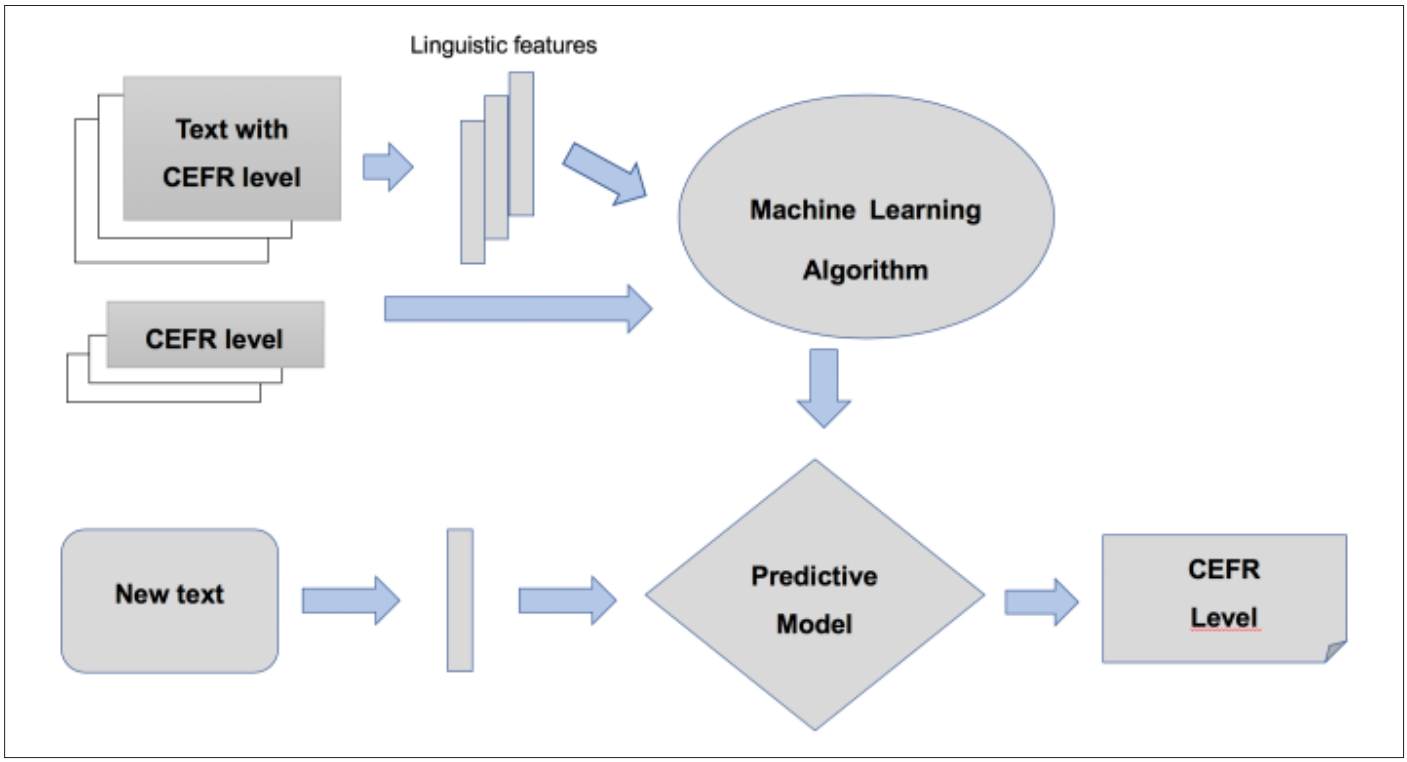


Figure 1. Supervised learning for CEFR-J RLDs.

The strength of this machine learning approach is in knowing the relative importance of the predictive features used for the classification. In our case, the question by which grammatical items play an important role in classification. By English Profile, these features are called 'Criterial Features' (Hawkins and Filipović 2012). A feature is *criterial* when the occurrences of this feature is so prominent at the given CEFR level that it helps distinguish that CEFR level from the rest. To prove this, we required information that this feature is significantly more frequent at a given CEFR level than others. To make matters more complicated, the CEFR level decision by humans is made not solely on a single feature but a bundle of lexical or grammatical features. Therefore, we used this machine learning algorithm not only to create a model to best predict the CEFR levels, but also to select the best combination of grammatical features as predictors.

To this end, we prepared two types of corpora, the ELT textbook corpus as 'input' and the learner corpus as 'output'. These two types of corpora were necessary in order to produce RLDs for both teaching and assessment purposes. The 'input' corpus is a collection of CEFR-based course books published in the UK. There are very few CEFR-based English textbooks (Naganuma et al. 2015) published in Japan, so course books published in the UK after the release of the CEFR in 2001 were collected and their content examined to see whether the textbooks were designed with appropriate CEFR levels in mind. In total, 96 textbooks were sampled. They were all scanned with an OCR and prepared in XML format. Each piece of textbook data in the corpus was tagged for CEFR level, section information for different skills (4 skills and grammar), part-of-speech and lemma for each word. The data set (c. 1,640,000 tokens) was prepared for both normal text processing and concordancing using Sketch Engine[9].

The 'output' corpus comprises two sets of learner corpora: the JEFLL Corpus (Tono 2007) and the NICT JLE Corpus (Izumi et al. 2004). The JEFLL Corpus is a collection of approximately 10,000 secondary school students' written compositions (size: 0.7 million), and the NICT JLE Corpus is a collection of oral interview test scripts by 1,280 test-takers (size: 2 million). Both sets of data were originally gathered without CEFR levels, but for this project all the sample texts were aligned to the CEFR levels.

9. http://www.sketchengine.co.uk

The extraction of grammar items from the two types of corpora was mainly conducted by my colleagues in the CEFR-J project (Ishii 2016; Ishii and Tono 2016). Altogether, approximately 500 grammar items were automatically extracted by using a set of pattern matching queries for each item. The frequencies and dispersion measures were obtained for each grammar category at all the CEFR levels and the matrix of [grammar category] x [each text with CEFR-levels] was used for machine learning. Several machine learning algorithms were tested, and random forest[10] and ranking Support Vector Machine (SVM)[11] were used for the final analysis (Tono 2017).

The CEFR-J Grammar Profile was released as a dataset first in March 2018[12], followed by the English teacher-friendly version in Fall 2018.

3.2.3 The CEFR-J Text Profile

Another important aspect of CEFR-level criteria is the characteristics of texts provided as input to learners at given CEFR levels. While a lot of readability measures have been proposed (cf. DuBay 2004), many of them have mainly been concerned with word levels and sentence length and have not included more complex lexical and syntactic features. The RLD project described above revealed more detailed vocabulary and grammar features relevant to each CEFR level. It is the co-occurrences of those linguistic features in a text that could serve as criteria for a particular CEFR level.

To this end, we extracted various textual features such as the CEFR levels of words in the text, the length of clauses and sentences, the number of verbs in the sentence, the depth of parsed tree of the sentence, and the ratio of difficult words in the noun phrases with more than two depth of trees. Currently, the profile information about the CEFR-level text characteristics is only available for written texts, but in the future, we hope to provide text features for spoken texts as well. For details, see Mizushima et al. (2016) and Uchida (2018).

4 Using the CEFR-J for other languages

So far, the historical development of the CEFR-J and its related language teaching and learning resources has been described in detail. Originally, the CEFR-J was designed to respond to the specific needs of English language teaching in Japan, but recently there is a growing interest in adopting the CEFR-J back into the CEFR itself or applying the framework developed for the CEFR-J to foreign languages other than English. For instance, in the Council of Europe (2017), they too added Pre-A1 level to the entire scale, as the CEFR-J originally proposed, and a large number of young learners' descriptors were supplied, for which approximately 30 descriptors were adopted from the CEFR-J.

Tokyo University of Foreign Studies (TUFS), where the author works, is the only national university in Japan that specialises in foreign language teaching with 28 foreign language majors. In 2014, TUFS launched a government-funded project called the Super Global University Program, where special focus is given to the development of a systematic program for teaching and assessment of the 28 foreign languages that TUFS students can major in. The university decided to use the CEFR-J as a core framework and I was appointed as the principal investigator of the CEFR-J x 28 project[13].

Table 3 shows the list of languages offered as majors at our institution:

10. Random forest is an ensemble learning method to build predictive models based on multiple decision trees (Breiman, 2001).
11. Ranking SVM is a variant of Support Vector Machine to deal with ranking data for classification. See Joachims (2002).
12. http://cefr-j.org/download.html#cefrj_grammar
13. The project used to be called the 'CEFR-J x 27', but recently one more language was added to the majors, thus now we have 28 language majors.

Table 3. *The list of languages for the CEFR-J x 28 project*

English	Japanese	German	French	Spanish
Cambodian	Russian	Chinese	Korean	Czech
Vietnamese	Thai	Urdu	Polish	Korean
Portuguese	Malay	Filipino	Turkish	Hindi
Mongolian	Laotian	Italian	Arabic	Persian
Indonesian	Burmese	Bengali		

5 A general approach for developing pedagogical resources

In the CEFR-J x 28 project, we share the CEFR-J as a common framework, which is also linked to the original CEFR as a foundation. The advantage of using the CEFR-J is its detailed sub-levels. There are four sub-levels up to A1 (Pre-A1, A1.1-1.3), followed by additional six levels from A2 to B2 (A2.1, A2.2, B1.1, B1.2, B2.1, B2.2). These levels almost correspond with the recently updated CEFR levels (Council of Europe, 2018). As was illustrated in the RLD work (Section 3), a set of resources such as the CEFR-J Wordlist, the CEFR-J Grammar Profile, and the CEFR-J Text Profile are available, which provided a good starting point for our project to explore the possibility of converting English resources into each language, using automatic techniques such as machine translation.

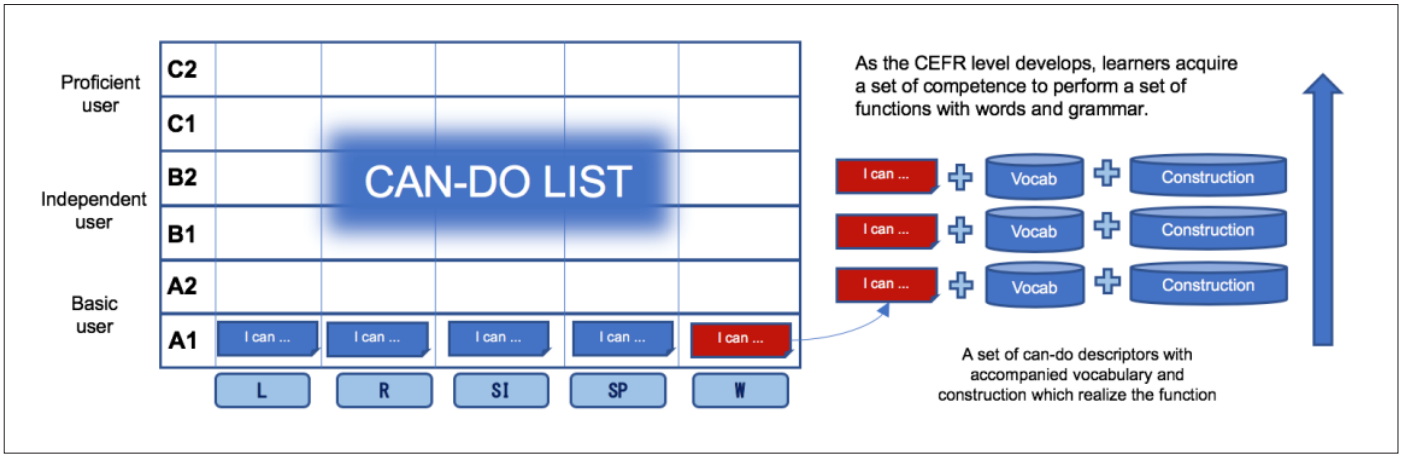


Figure 2. The relation between a set of Can Do descriptors and lexical and grammatical resources.

Figure 2 shows our basic approach. Before converting the English resources into 27 other languages, the level at which automatic conversion should be attempted, required careful consideration. If a simple one-to-one machine translation was made for a certain word in English, the chances are that most content words (nouns and adjectives) with a single meaning can be converted fairly accurately into a given language, whereas most of the grammatical words and polysemous words will fail, due to various structural and semantic mismatches between the two languages.

However, consider the level of language functions such as "express likes or dislikes." A set of model constructions can be selected to realize such functions, such as "I like ...", "I don't like ...", "Do you like ...?" or "What do you like?" At this level, translating English phrases into the counterpart in a given language is more likely to be successful, due to the availability of contextual information derived from specified language functions. Also, if specific content words, e.g. sports, food, favourite pastimes, are used with these constructions to form a sentence, then the automatic translation of these sentences is more likely to succeed, given the detailed context provided at a sentence level.

Interestingly, the CEFR provides this very list of Can Do descriptors for each level. Therefore, we have decided to first compile a list of words and constructions that should go with each set of Can Do descriptors at a given CEFR-J level. This resource is called the CEFR-J Can Do Phrase Database. This phrase database serves as the primary input to feed into a machine translation system. For the first test run, we used Google Translate. In the past few years, the level of machine translation has drastically improved since the innovation made by neural machine translation (NMT). The translation quality of Google Translate has become impressively high, compared to a few years ago.

Table 4 shows some examples of the CEFR-J Can Do Phrase Database and its multilingual version.

Table 4. ***Sample database entries for CEFR-J: A1.2 spoken interaction descriptor***

CEFR-J A1.2 spoken interaction Can Do	I can exchange simple opinions about very familiar topics such as likes and dislikes for sports, foods, etc., using a limited repertoire of expressions, provided that people speak clearly.
Function	Expressing pleasure, liking
Construction	I like + NP (very much).
Japanese	NP を(とても) 好きです
Arabic	أنا أحب + NP
Turkish	NP + (çok) severim.
Thai	ฉันชอบ + NP (มาก ๆ)
Malays	Saya suka + NP sangat
Burmese	NPကို အရမ္းကိုက္ၾတယ္။
Indonesian	Saya suka + (sekali)
Bengali	আমি + NP খুব পছন্দ করি).
Chinese	我（非常）喜欢+NP
German	Ich mag + NP (sehr gerne).
Mongolian	Маш их
Russian	Мне (очень) нравится

We are now at a preliminary stage, evaluating the output of machine translation over various types of resources, including the CEFR-J Wordlist itself as well as a part of the Phrase Database. A team of linguists, computer engineers, as well as language instructors work together to make the most of the CEFR-J and its related resources for creating language teaching and learning resources for 27 other languages (Fig. 3).

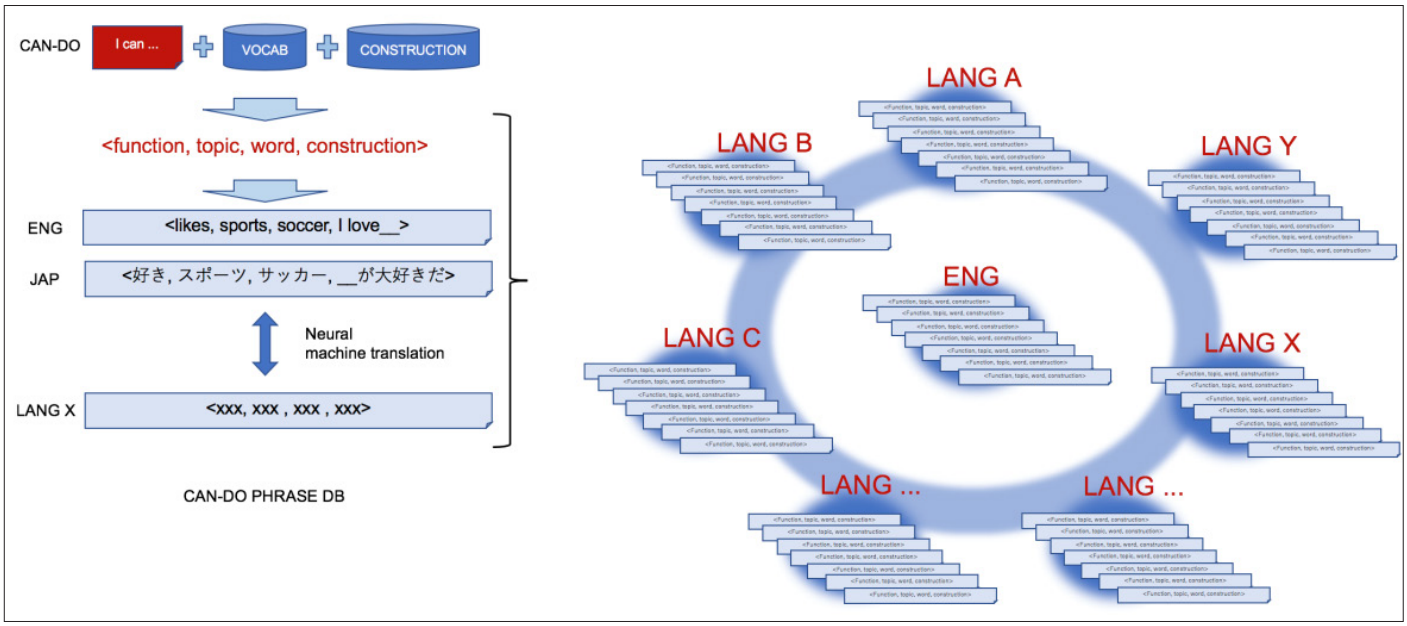


Figure 3. The image of CEFR-J-based pedagogical resources shared among 28 languages.

6 Developing e-learning tools and apps for teaching 28 languages

As we develop the CEFR-J pedagogical resources for 28 languages, three types of e-learning tools and applications have been developed.[14]

6.1 The Flash Card Vocab Builder

An Apple/Android app for learning vocabulary in 28 different languages called the Flash Card Vocab Builder (FCVB) was developed. This is a simple flash card type application, in which learners can choose any one of 28 languages and learn content words such as verbs, nouns and adjectives. One unique feature is that the words are grouped together according to the thematic categories based on *Threshold Level* (van Ek and Trim 1990) as well as the CEFR levels determined by English equivalents. In this way, they can learn basic everyday vocabulary in a given language using flash cards on their smartphones (Fig 4).

Figure 4. The Flash Card Vocab Builder: (a) Language menu, (b) CEFR levels and (c) Themes.

14. Currently, these tools and apps are available for internal use only. TUFS has a plan to make them open to public once the SGU project is over.

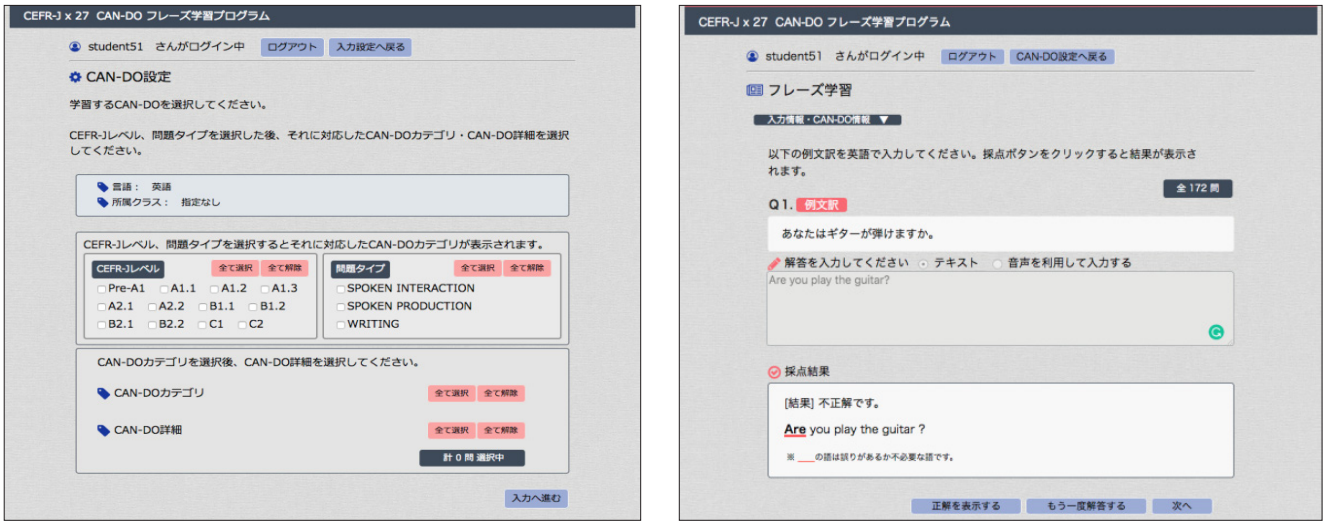
On the menu, you can select one of 28 languages. Once you select a language, you will be asked to choose a CEFR level you want to study, which will take you to the list of words grouped together according to the thematic domains in that specific Threshold Level. The translation can be displayed in either English or Japanese, so this app can be used for speakers whose L1 is one of the 27 languages and want to study Japanese.

Figure 5 shows the main study page. You can see the card in the centre, and you just flip the page to the left (Don't know yet) or to the right (I got it!). The log file is kept on the server and teachers can check each learner's progress in terms of how many words have been learnt for each CEFR level and in which thematic categories.

Figure 5. The main study page of the FCVB.

6.2 The Can Do Sentence Builder

The second tool is a web writing tutor. Figure 6-(a) shows the menu of specific CEFR levels and skills. When a learner chooses levels and skills, specific Can Do descriptors will be displayed. When you select particular descriptors, you will be taken to a writing practice screen shown in Figure 6-(b).



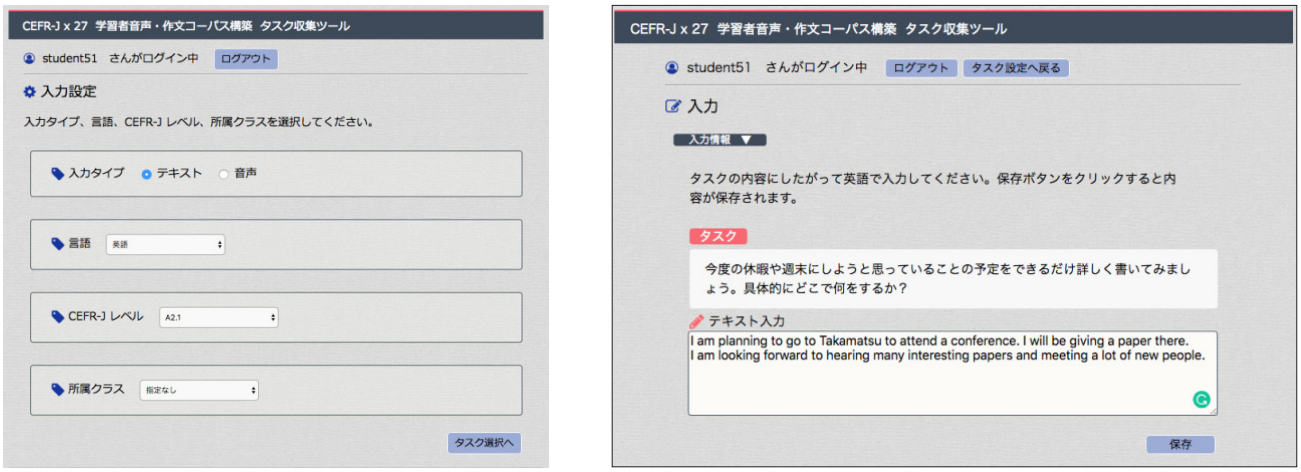
(a) (b)

Figure 6. The Can Do Sentence Builder.

The sentence cues will be provided in either Japanese or English. You translate the sentence into the target language. In this case, "Can you play the guitar?" is the target sentence. Any character strings that do not match the target will be highlighted as in the bottom of the screen, suggesting either something is missing (omission error), something is redundant (addition error) or some forms are wrong (misformation error). This judgement is based on the comparison between the target string and the input string only as the current version cannot deal with multiple possibilities of translations yet, But at least if you have specific Can Do descriptors and their functions, it would be useful to go through basic sentences comprised of useful constructions and topic vocabulary. The nice thing about this tool is that all the 28 languages have the same format. Once you learn one language, it is possible to learn additional language in the same way, or even in parallel.

6.3 The Can Do Task-Based Spoken/Written Corpus Collection Tool

The final tool is a web-based corpus collection interface. At this site, students can choose from the main menu a choice of their language and their estimated CEFR levels, and they will be shown a list of topics for speaking or writing, tuned to a particular CEFR level selected, as in Figure 8-(a).



(a) (b)

Figure 8. The Can Do Task-Based Spoken/Written Corpus Collection Tool.

Then students will be taken to the work space, shown in Figure 8-(b), where the essay task based on the Can Do descriptor is displayed and they are asked to write their essays in the field at the bottom. When they click on the "save" button, the whole essay data, together with all the person- and task-related metadata, will be saved onto the server. The same thing can be done for speaking tasks, where students press the record button and speak using the built-in microphone. In the current system, English and Chinese can be processed using a voice recognition system[15], which will automatically convert your speech into orthographical data.

This is a quite simple design, but if used properly, it would be a very useful tool to collect learner production data in a very cost-efficient way. One can assign either spoken or written tasks related to target Can Do descriptors and ask students to record their performance online. If designed properly, the system would be useful in collecting texts for different text types and stylistic variations across languages, which would be quite useful to cross-compare the effects of tasks on the definition of spoken and written production. It is also possible to keep track of students' progress if a series of spoken or written output is recorded on the server during the course. The system saves all the speech and text

15. For this, Sinewave Inc. provides technical support on our system.

data for individual learners with all the details of task and student information. This system can be used for both teaching and research. In the classroom, teachers can provide more valid CEFR-based grading by evaluating students' performance in both speaking and writing with this system. The system can gather all the students' data in different languages from the beginning of their study to when they leave university. It can contribute to the creation of L2 learners' production data in multiple languages and this has much potential for future research as big data.

7 Conclusion

With the growing influence of the CEFR, attempts have been made to reconstruct the entire framework of teaching and assessing foreign languages using the CEFR. The CEFR-J Project is one such example. This study has reported ongoing projects applying CEFR-J resources for teaching different languages. While criticism still persists about the validity of the CEFR as a generic language framework, the present author believes that the validation process of such a framework and accompanying resources are quite intriguing as a research topic. The evaluation of our multilingual resource development based on the CEFR-J is yet to be seen, but the approach taken by the CEFR-J x 28 project is moving in a promising direction in that resource-rich languages such as English could give support to under-resourced languages in terms of language teaching and learning content and methods.

8 Acknowledgements

This work is supported by JSPS KAKENHI Grant Number JP16H01935 and JP18H03658.

9 References

Breiman, Leo. 2001. Random Forests. *Machine Learning* 45(1): 5-32.

Council of Europe. 2001. *The Common European Framework of Reference for Languages: Learning, teaching, assessment.* Strasbourg: Council of Europe.

Council of Europe. 2018. *Common European Framework of Reference for Languages: Learning, teaching, assessment. Companion Volume with New Descriptors.* February 2018.

DuBay, William H. 2004. *The Principles of Readability*. 2004.

Hawkins, John A. & Luna Filipović. 2012. *Criterial Features in L2 English: Specifying the Reference Levels of the Common European Framework*. English Profile Studies, Vol 1. Cambridge: Cambridge University Press.

Ishii, Yasutake. 2016. Grammatical Items for Creating the CEFR-J Grammar Profile. Paper presented at the Symposium on the CEFR-J RLD Project: Developing Grammar, Text and Error Profiles Using Textbook & Learner Corpora. March 5, 2016. Tokyo University of Foreign Studies.

Ishii, Yasutake & Yukio Tono. 2016. CEFR-J Grammar Profile no tame no Bunpo koumoku hindo-chosa. [A frequency analysis of grammar categories for the CEFR-J Grammar Profile.] Paper presented at the Symposium on the CEFR-J RLD Project: Developing Grammar, Text and Error Profiles Using Textbook & Learner Corpora. March 5, 2016. Tokyo University of Foreign Studies.

Izumi, Emi, Kiyotaka Uchimoto & Hitoshi Isahara. 2004. The NICT JLE corpus: Exploiting the language learners' speech database for research and education. *International Journal of the Computer, the Internet and Management, 12*(2), 119–125.

Jin, Yan, Zunmin Wu, Charles Alderson & Weiwei Song. 2017. Developing the China Standards of English: challenges at macropolitical and micropolitical levels. *Language Testing in Asia, 7*:1. https://doi.org/10.1186/s40468-017-0032-5. (accessed 02 October, 2018.)

Joachims, Thorsten 2002. Optimizing the search engine using clickthrough data. *Proceedings of the ACM Conference on Knowledge Discovery and Data Mining* (KDD), ACM, 2002.

Mizushima, Kohei, Yuki Arase & Satoru Uchida. 2016. CEFR junkyo-kyozai ni okeru goi bunpo no tokucho bunseki to level jido bunrui. [Lexico-grammatical feature analysis and automatic level classification using CEFR-based teaching materials.] *Journal of Natural Language Processing* 2016: 789-792.

Naganuma, Naoyuki, Noriko Nagai & Fergus O'Dwyer. 2016. *Connections to thinking in English: The CEFR-informed EAP textbook series* (B1 to B1). Tokyo: Asahi Press.

Negishi, Masashi, Tomoko Takada & Yukio Tono. 2013. A progress report on the development of the CEFR-J. In E.D. Galaczi & C. J. Weir (Eds.), *Exploring language frameworks, Proceedings of the ALTE Krakow Conference*, July 2001, pp. 135-163.

Negishi, Masashi & Yukio Tono. 2016. An update on the CEFR-J project and its impact on English language education in Japan. *Studies in Language Testing*, 44: 113-133.

North, Brian, Ortega, Angeles & Susan Sheehan. 2010. *The British Council-EAQUALS Core Inventory for General English.* Accessed via https://englishagenda.britishcouncil.org/ sites/default/files/ attachments/books-british-council-eaquals-core-inventory.pdf. (accessed 28 September, 2018.)

Tono, Yukio (Ed.). 2007. *Chukousei 1-man nin no eigo corpus: The JEFLL corpus* [A corpus of 10,000 Japanese secondary school students' writings: the JEFLL corpus]. Tokyo: Shogakukan.

Tono, Yukio (Ed.). 2013. *The CEFR-J handbook*. Tokyo: Taishukan Shoten.

Tono, Yukio. 2017. The CEFR-J and its impact on English language teaching in Japan. *JACET international convention selected papers*, 4, 31–52.

Schmidt, Helmut. 1994. Probabilistic Part-of-Speech Tagging Using Decision Trees. *Proceedings of International Conference on New Methods in Language Processing*, Manchester, UK.

Uchida, Satoru. 2018. CEFR-J Text Profile no Katuyou. [How to use the CEFR-J Text Profile] A talk given at the CEFR-J 2018 Symposium, Seijo University. March 18, 2018.

van Ek, Jan A. & John Trim. 1991. *Threshold Level 1990*. Cambridge: Cambridge University Press.

10 Biography

Yukio Tono is a professor at Tokyo University of Foreign Studies. His primary research interests include corpus linguistics, L2 vocabulary acquisition, L2 lexicography and the application of corpora in language learning and teaching. He is a research group project leader of the CEFR-J, and his current research focuses on how corpora can be used to identify criterial features for the CEFR-J.