CEFR Journal

Research and Practice

CEFR JOURNAL—RESEARCH AND PRACTICE

Table of Contents

journal@cefrjapan.net

Mission statement

The CEFR Journal is an online, open-access, peer-to-peer journal for practitioners and researchers. Our editorial advisory board comprises comprises stakeholders on a wide range of levels and from around the world. One aim of our journal is to create an open space for exchanging ideas on classroom practice and implementation related to the CEFR and/or other language frameworks, as well as sharing research findings and results on learning, teaching, and assessment-related topics. We are committed to a strong bottom-up approach and the free exchange of ideas. A journal by the people on the ground for the people on the ground with a strong commitment to extensive research and academic rigor. Learning and teaching languages in the 21st century, accommodating the 21st century learner and teacher. All contributions have undergone multiple double-blind peer reviews. We encourage you to submit your texts and volunteer yourself for reviewing. Thanks a million. *<journal@cefrjapan.net>*

Aims, goals, and purposes

Our aim is to take a fresh look at the CEFR and other language frameworks from both a practitioner's **and** a researcher's perspective. We want the journal to be a platform for all to share best practice examples and ideas, as well as research. It should be globally accessible to the wider interested public, which is why we opted for an open online journal format.

The impact of the CEFR and now the CEFR Companion Volume (CEFR/CV) has been growing to previously wholly unforeseeable levels. Especially in Asia, there are several large-scale cases of adoption and adaptation of the CEFR to the needs and requirements on the ground. Such contexts often focus majorly on English language learning and teaching. However, there are other language frameworks, such as the ACTFL and the Canadian benchmarks, while the Chinese Standard of English (CSE) is also on its way. On the one hand there is a growing need for best practice examples in the form of case studies, and on the other hand practitioners are increasingly wanting to exchange their experiences and know-how. Our goal is to close the gap between research and practice in foreign language education related to the CEFR, CEFR/CV, and other language frameworks. Together, we hope to help address the challenges of 21st century foreign language learning and teaching on a global stage. In Europe, many take the CEFR and its implementation for granted, and not everyone reflects on its potential uses and benefits. Others are asking for case studies showing the effectiveness of the CEFR and the reality of its usage in everyday classroom teaching. In particular, large-scale implementation studies simply do not exist. Even in Europe, there is a center and a periphery of readiness for CEFR implementation. It is difficult to bring together the huge number of ongoing projects from the Council of Europe (CoE), the European Centre for Modern Languages (ECML), and the EU aiming to aid the implementation of the CEFR. This results in a perceived absence in the substance of research. Outside Europe, the CEFR has been met with very different reactions and speeds of adaptation and implementation. Over the last few years, especially in Asia, the demand by teachers for reliable (case) studies has been growing.

For more than a decade, the people behind this journal – the Japan Association for Language Teaching (JALT) CEFR & Language Portfolio special interest group (CEFR & LP SIG) – have been working on a number of collaborative research projects, yielding several books and textbooks, as well as numerous newsletters. This is a not-for-profit initiative; there are no institutional ties or restraints in place. The journal aims to cooperate internationally with other individuals and/or peer groups of practitioners/researchers with similar interests. We intend to create an encouraging environment for professional, standard-oriented practice and state-of-the-art foreign language teaching and research, adapted to a variety of contexts.

Editorial advisory board

- Brian North (CEFR co-author),
- Nick Saville (Cambridge Assessment English, UK),
- Yukio Tono (Tokyo University of Foreign Studies, Japan),
- Neus Figueras (Generalitat de Catalunya/University of Barcelona, Spain),
- Bärbel Kühn (TU Darmstadt, Germany),
- Fergus O'Dwyer (University College Dublin, Ireland),
- Pham Thi Hong Nhung (University of Foreign Languages, Hue University, Vietnam),
- Marita Härmälä (Finnish Education Evaluation Centre, Finland),
- Vincent Folny (CIEP, France),
- Cristina Rodríguez (EOI Santiago de Compostela, Spain),
- Irina Pavlovskaya (St. Petersburg State University, Russia),
- Grant Black (Black Inc. Consulting, Japan)

Journal editorial team

- Morten Hunke (g.a.s.t. | TestDaF-Institut, Germany) **– *Editor-in-Chief Volume 1***,
- Maria Gabriela Schmidt (Nihon University, Tokyo, Japan),
- Alexander Imig (Chukyo University, Nagoya, Japan)

Editing and proofreading team

- Emma Craig, Amy Curry, Susanne Eriksson, Juhi Gupta, Bettina Hermoso-Gomez, Laura Elizabeth MacNay, James Mendelsohn, Jael Räker, Hannah Joanna Sheeran, Michael Stout

Layout

- Malcolm Swanson

Submission (Call for papers)

This journal attempts to fall somewhere in between an inaccessible academic journal (long waiting times, fairly strict guidelines/criteria) and a newsletter (practical in nature but lacking in theoretical support/ foundation), linking research of a practical nature with relevant research related to foreign language education, the CEFR, other language frameworks, and the European Language Portfolio. While the CEFR was introduced by the Council of Europe and intended for use, first and foremost, within Europe, the influence of the CEFR now has to be attested in many places beyond European borders. It has become a global framework, impacting a variety of aspects of language learning, teaching, and assessment across countries and continents beyond the context for which it was originally created. As such, there is a pressing need to create a quality forum for sharing research, experiences, and lessons learned from applying the CEFR in different contexts. This journal provides such a forum where people involved or interested in processes of applying the CEFR can share and learn from one another.

We are continuously seeking contributions related to foreign language education, the CEFR, other language frameworks, and the European Language Portfolio. We are particularly interested in specific contextual adaptations.

Please contact the editors and submit to:

journal@cefrjapan.net

Editorial

Morten Hunke

This is the maiden issue of our new **CEFR Journal – Research and Practice**. It has taken us a little longer to publish than we had expected and hoped for, but we are glad to be able to finally introduce our brand-new online journal to the public. We envisage this journal as an accessible platform for different kinds of learning, teaching and research activities in the field of the CEFR, language frameworks, and portfolios. Reports on best practice and work in progress are equally as welcome as article/book reviews and academic articles. It goes without saying that the journal stands firmly on the grounds of due diligence and quality assurance. All submissions undergo double-blind peer reviews by at least two reviewers.

In this first issue, we are proud to present to you an illustrious collection of texts from around the globe. We kick off by exploring some of the after-effects of the extensive CEFR-J project in Japan. In this progress report, the reader is presented with glimpses of how such a huge project outside Europe now contributes to language learning and teaching resources globally as can be seen in the CEFR-Jx28 project in *Coming full circle: From CEFR to CEFR-J and back* (Yukio Tono, Tokyo University of Foreign Studies, Japan).

This opening article is followed by some meticulous bibliometric research on the width and breadth of scholarly work relating to the CEFR by Judith Runnels (University of Bedfordshire, UK) and Vivien Runnels (University of Ottawa, Canada) in *Impact of the Common European Framework of Reference: A bibliometric analysis of research from 1990-2017*.

Next up, Irina Pavlovskaya and Olga Lankina from the University of St. Petersburg (Russia) showcase early use of the newly added feature of mediation in the publication of the CEFR Companion Volume (CEFR/CV): *How new CEFR mediation descriptors can help to assess the discussion skills of management students – global and analytical scales*.

Vietnam makes for another astounding example of adoption and adaptation of the CEFR in an Asian country in order to achieve massive nationwide changes to the entirety of (English) language teaching from school to higher education level. How such drastic alterations affect teachers having to conform to the new system is described in an article by Pham Thi Hong Nhung and Le Thi Thanh Hai (Hue University of Foreign Languages, Vietnam): *Implementing the CEFR at a Vietnamese university: General English language teachers' perceptions*.

In *Jumping through hoops and keeping the human-in-the-loop*, Maria Gabriela Schmidt and I myself had the opportunity to interview Dr. Nick Saville (Director of Research and Thought Leadership at Cambridge Assessment English, UK) at the JALT International Conference in Tsukuba in November 2017. This interview looks at both the history of the CEFR in Japan as well as issues surrounding language testing and the role of artificial intelligence in the sphere. It also offers some insights into the background history of the JALT CEFR & LP SIG and it helps to contextualize how this journal came into existence.

Thus, is the maiden issue of the **CEFR Journal** rounded off. It has been an honor to serve as the Editor-in-Chief for this first issue. I would like to express my deepest gratitude to everyone who made this journal possible: the authors, the reviewers, the proofreaders, the editorial advisory board, and especially the JALT CEFR & LP SIG officers and members. We have been working together closely for more than 10 years and have realized a fair number of CEFR-related projects. May this journal gradually become the platform of mutual support and stimulus to foreign language professionals around the world that we are envisioning it to be.

—Bochum (Germany) & Tokyo (Japan), May 2019

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

Yukio Tono, Tokyo University of Foreign Studies

The CEFR-J project was launched in Japan in 2008. The CEFR-J gives a set of Can Do descriptors for 10 CEFR sublevels (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.

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
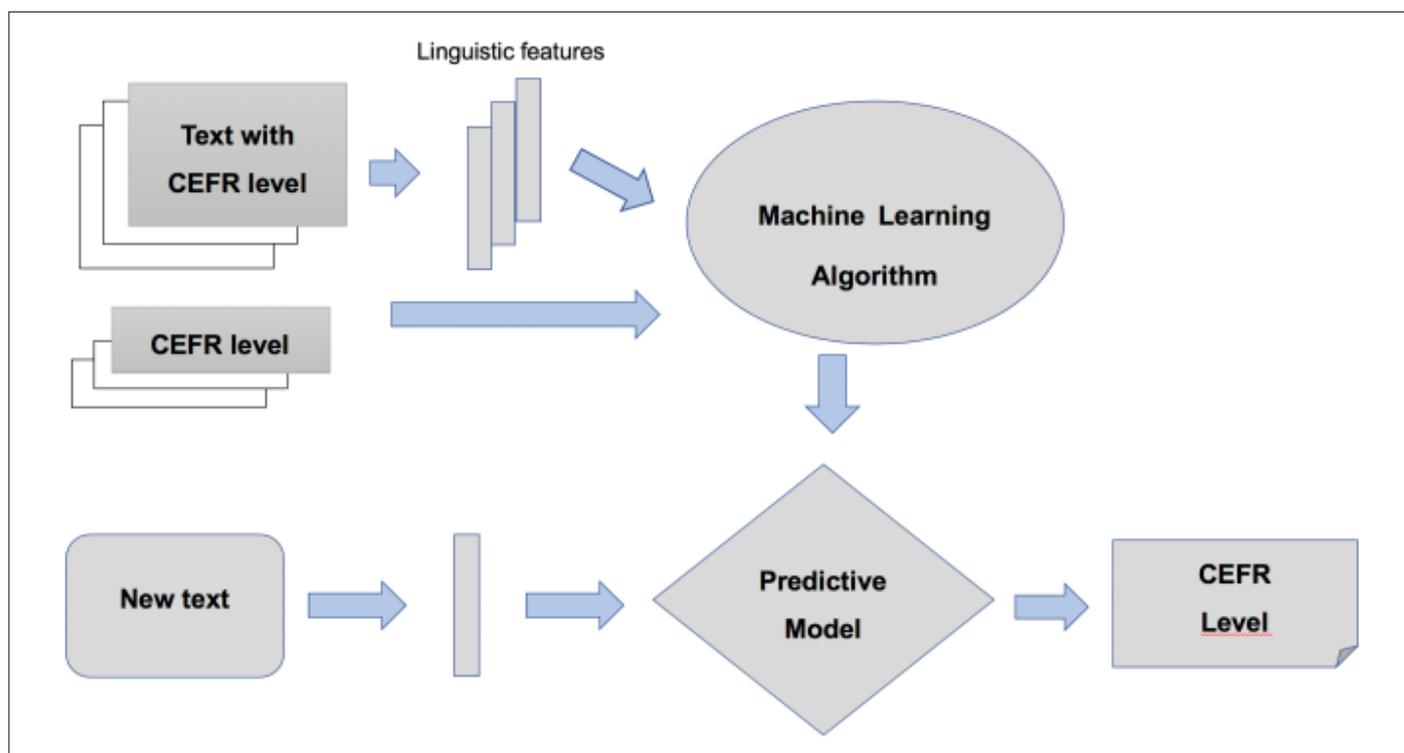
Figure 1. Supervised learning for CEFR-J RLDs.

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	NP + أنا احب
Turkish	NP + (çok) severim.
Thai	ฉันชอบ + NP (มาก ๆ)
Malays	Saya suka + NP sangat
Burmese	NPကို အရမ်းကြိုက်တယ်။
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).
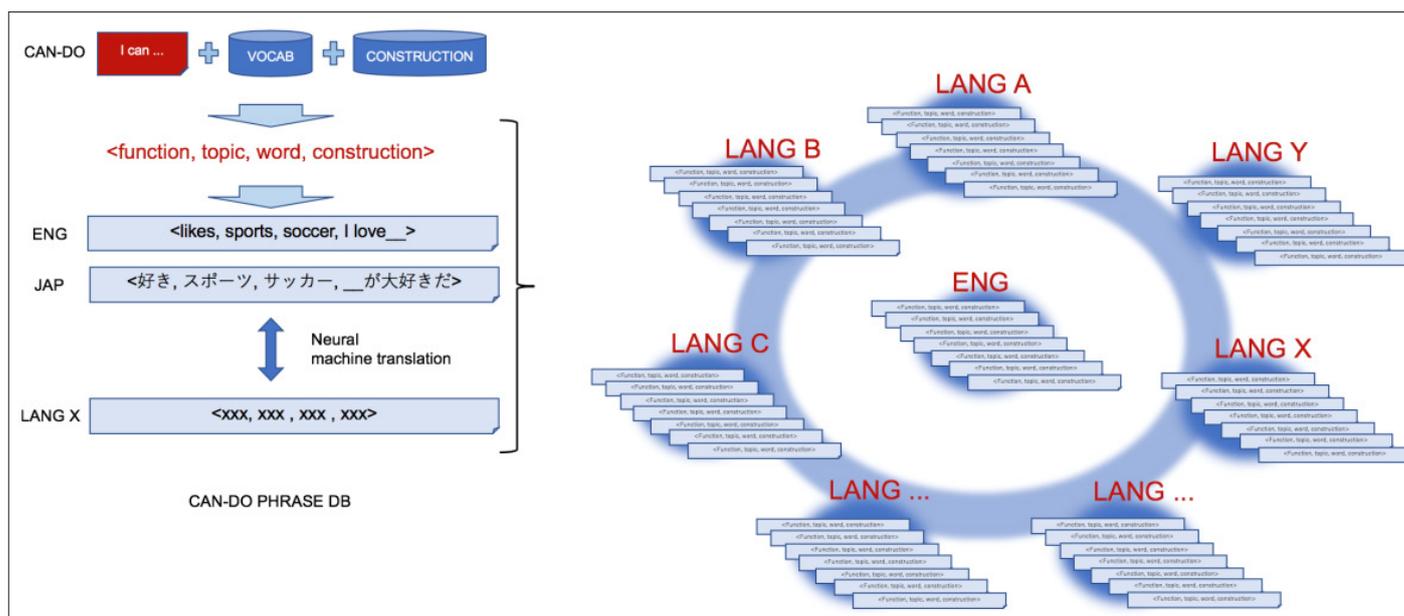
Figure 3. The image of CEFR-J-based pedagogical resources shared among 28 languages.

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).
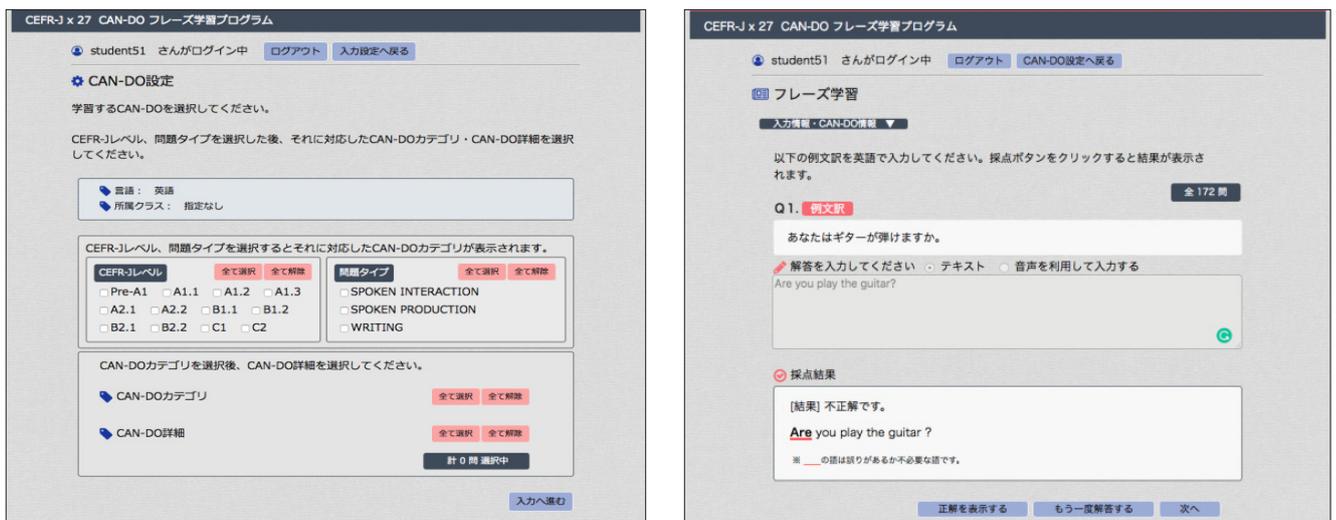
(a)   (b)

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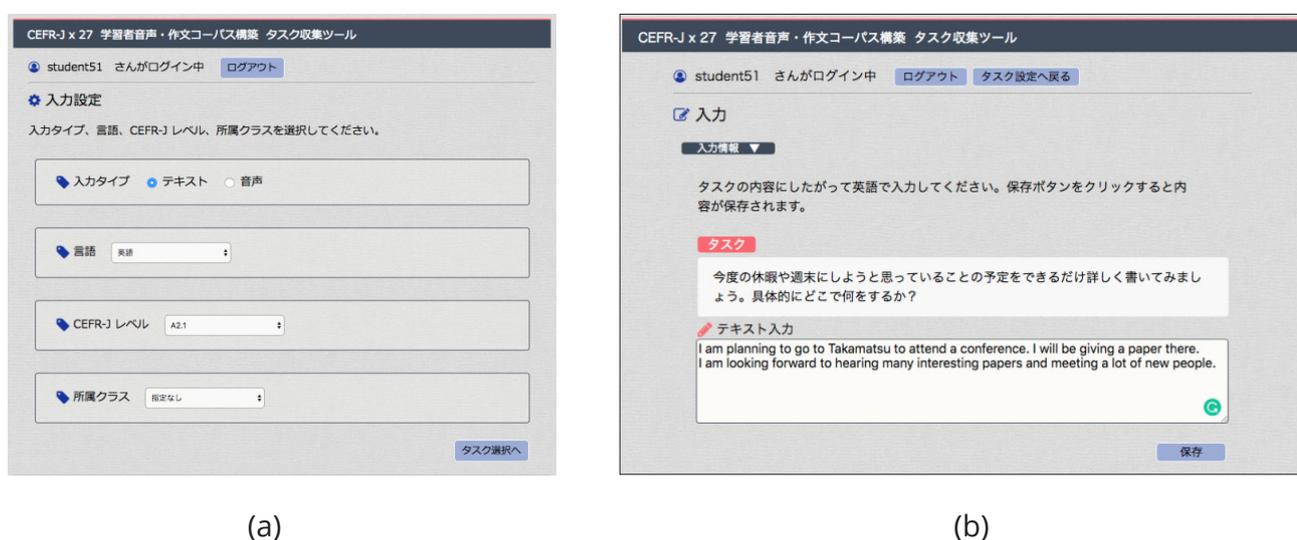
Figure 8. The Can Do Task-Based Spoken/Written Corpus Collection Tool.

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

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.

Impact of the Common European Framework of Reference—A bibliometric analysis of research from 1990-2017

Judith Runnels, University of Bedfordshire

Vivien Runnels, University of Ottawa

Published in 2001, the Common European Framework of Reference for Languages (CEFR), a reference framework which informs teaching, learning and assessment in language education, appears to be increasingly recognized, referenced and utilized in language education contexts worldwide. To date however, the extent, provenance and adoption of the collected body of knowledge concerning the CEFR has yet to be systematically analysed, rendering it difficult for any conclusions to be made about its impact. A bibliometric analysis was therefore conducted to explore the CEFR from the document's more formal origins in 1990 to the end of 2017 for the bibliometric indicators of number of publications per year, geographical location of research, highly cited works and journals with the highest number of relevant publications. The findings show that research on the CEFR has increased significantly over the examined time. The majority of publications with a focus on the CEFR are European, but numbers are increasing in geographical areas outside of Europe, and particularly in Asia. The framework is discussed in numerous types of publications covering a range of topics in language education. These findings suggest that the CEFR has been used in contexts beyond its origins and has influenced many aspects of language education around the globe. Diffusion of innovations theory suggests that the CEFR's impact and influence is likely to increase over the next ten years in and outside of Europe and especially in Asia.

Keywords: CEFR, bibliometric analysis, bibliometric indicators, adoption, diffusion, diffusion of innovations, educational innovation

1 Introduction

The Common European Framework of Reference (CEFR) is the culmination of decades of work from a number of participating institutions and contributors in Europe, designed to improve the communication and mutual understanding of language education stakeholders on the topics of language learning, teaching, and assessment in all European languages (Council of Europe 2001). The CEFR is also a policy tool based on the tenets that education is a human right, and that multilingualism and plurilingualism can increase mutual understanding among individuals with different linguistic and cultural backgrounds, thus building inclusive societies (Council of Europe 2001; 2018). According to the CEFR, a plurilingual approach to language education is one that recognizes the interrelationships and interactions between language and culture and that communicative competence is built according to these interactions. This means that an individual "can call flexibly upon different parts of this competence to achieve effective communication with a particular interlocutor" (Council of Europe 2001: 5). The plurilingual approach emphasizes that as an individual person's experience of language in its cultural contexts increases, from the language of the home to that of society, and then to the languages of other peoples (whether learnt at school or college, or by direct experience), he or she does not keep these languages and cultures in strictly separated mental

compartments. Rather, the person builds up a communicative competence to which all knowledge and experience of language contributes, and in which languages interrelate and interact.

The CEFR was more formally conceived at the *Transparency and Coherence in Language learning in Europe: objectives, evaluation, certification* Symposium, held in Switzerland in 1991 (Council of Europe 2001b). In 1995, a draft of the framework was produced for evaluation with further revisions resulting in the first version being published in English and French in 2001. Used all over the world, it is now available in 40 languages with a companion document published in 2018 providing recently updated descriptors (Council of Europe 2018b). Many scholars refer to its success and increasing popularity (Alderson 2007; Carty 2014; Council of Europe 2005; Figueras 2012; Li and Zhang 2004; Martyniuk and Noijons 2007; Nagai and O'Dwyer 2011; O'Dwyer 2014; O'Dwyer et al. 2017; Papageorgiou 2014; Valax 2011). Furthermore, the CEFR is identified as having had a positive impact in a number of domains in language education, such as curriculum design and development, pedagogy and teacher education (Little 2006; Hulstijn et al. 2010; Faez et al.2012; Jones and Saville 2009; Little 2007; Figueras 2012; Piccardo et al. 2013; Eckes et al. 2005; Schärer 2007).

A handful of studies have explored the usage of the CEFR on an international level. For example, Valax (2011) considers how language teachers perceive the impact of the CEFR on curriculum design in two countries from each of the European, Asian and Oceanian continents. The Council of Europe surveys in 2005 and 2007 also looked at utilization of the CEFR in Europe and beyond (Martyniuk and Noijons 2007). Other studies have considered the CEFR's usage at national levels in countries such as Japan, Colombia and Vietnam (de Mejía 2011; Ngo 2017; Schmidt et al. 2017). However, the sampling of respondents in these works are rather limited and each focuses on vastly different aspects or users of the CEFR, which makes it difficult to generalize utilization of the CEFR in assessing its impact. To date, there has been little in the way of systematic analysis of the applied and theoretical body of literature on the CEFR. An examination of this literature could provide insight into the progression of research on the CEFR since its more formal conception around 1990 to 1991 and an exploration of its uptake or adoption and current impact.

1.1 Bibliometric analysis

One methodology to derive evidence for research profiling is a bibliometric analysis (Kostoff et al. 2001; Porter et al. 2002). Bibliometric analysis refers to methods used to assess a field of research through the examination of large-scale publication metadata (Borgman and Furner 2002; Xian and Madhavan 2014). It entails the quantifiable study of a body of literature to uncover historical development, patterns in publications or authorship, and usage over time (Tricco et al. 2008). Bibliometric analyses can provide a macro focus on a specific subject from a field of research, by incorporating a large range of works into numerical and graphical depictions of the field, in contrast to solely textual discussions summarizing content typically seen in some types of literature review (Porter et al. 2002). Such analyses can produce quantifiable estimates of productivity, importance, or visibility of research, can explore the occurrence of specific events within the literature (Koskinen et al. 2008), or can highlight collaborations between scientists in the field (Glänzel et al. 1999).

1.2 Focus of the study

To our knowledge, bibliometric analyses have not been widely utilized in language education, and certainly not to carry out a review of research on the CEFR. In this study we aim to explore the impact of the CEFR through an examination of the body of scholarly research related to it and its changes over time. 'Impact' is being used herein to refer to having a marked effect or influence. It does not refer to having a positive or negative impact on language education within the context where it was researched – it simply refers to the change over time in bibliometric indicators (either increases or decreases). Bibliometric indicators that reflect the extent (number of publications and number of publications per

year) and provenance of work (the source and geographical location of the publications and the most highly cited works) were thusly profiled (Van Leeuwen 2006). The implications these have on the CEFR's adoption and impact is considered. Such knowledge will not only allow for a better understanding of the characteristics or patterns in previous work performed on the CEFR, but may also suggest direction for future research in the field and inform policy and decision-making (Hanney et al. 2003, Mays et al. 2005, Milat et al. 2011, Koskinen et al. 2008, Van Leeuwen 2006).

2 Methods

An approach was employed that is commonly used in bibliometric analyses on emerging literatures similar to those described in Karakaya et al. (2014) and Koskinen et al. (2008). The five-step process involved the selection of i) literature search instruments, ii) a search term(s), iii) bibliometric indices, iv) the search itself, and v) the analysis of the search results.

2.1 Instruments

Glänzel et al.'s (1999) factors for the selection of a data source for a bibliometric analysis guided the decision to use Google Scholar and EBSCO Host as the literature search instruments. These factors include multidisciplinarity (which refers to the span of disciplines included), selectiveness (which refers to the criteria for inclusion – for instance, whether a publication is peer-reviewed or not), coverage (the extent to which it includes a record of all papers published in the discipline), and completeness (the extent to which information for each citation is complete).

Google Scholar is a publicly accessible web search engine that includes peer-reviewed papers, theses and dissertations, books, abstracts, articles from academic publishers, professional societies, universities, and other scholarly organizations (University of Wisconsin–Milwaukee 2014; Vine 2006). It is also compatible with free, publicly accessible software for performing bibliometric analyses called Publish or Perish (Harzing 2007). This program retrieves and analyses academic works from a number of databases and presents bibliometric statistics such as the number of citations, citations per year, and citations per author (Harzing 2007). EBSCO Host is an indexing engine that provides research databases tailored to the needs of libraries, corporations, or military institutions (EBSCO Industries 2016). Google Scholar was selected because of its accessibility and comprehensive coverage in social science (Harzing and Alakangas 2016) while EBSCO Host was selected because of its advanced sort and filter features and more detailed publication metadata, which allowed for the assessment of bibliometric indicators that could not have been assessed using Google Scholar alone.

2.2 Procedure

The search term 'Common European Framework of Reference' was selected for the bibliometric analysis due to having the highest number of hits on both databases when compared to a number of other terms that were pilot-tested (these included Common European Framework, Common European Framework of Reference, Common European Framework of Reference for languages, CEFR, and CEF). This term also resulted in a far higher number of relevant retrievals, and few false hits in comparison to the other keywords.

The bibliometric indicators used in the current study were selected because they provide estimates of overall productivity, productivity per year, important and impactful works, as well as a general understanding of where research is being conducted (Van Leeuwen 2006; Fagerberg 2009):

   i.   Number of publications
   ii. Number of publications per year
   iii. Source

iv. Most cited works

v. Geographical location

According to the information provided by each database, EBSCO Host and Google Scholar were both used for indicators i) and ii), EBSCO Host alone was used for iii) and v) and Google Scholar alone was used for iv).

2.3 Screening procedure

Using the keyword 'Common European Framework of Reference', a literature search from 1990-2017 was conducted in both EBSCO Host and Google Scholar. Each search was repeated (once in the morning and once in the afternoon) on two different days within the first week of 2018, although the same number of hits were obtained in each database each time.

Prior to recording the data, the resulting hits from the literature searches were screened for irrelevant literature. The first 1000 hits on Google Scholar by way of Publish or Perish (PoP) contained two articles that were not in reference to the CEFR. These articles were removed prior to any data recording or analysis. In EBSCO Host, non-print, audio, trade publications, and news sources were removed and manual verification of the first 500 remaining search hits confirmed that they all referred to the CEFR.

2.4 Number of publications and publications per year

Following the screening procedure, the total number of search hits was recorded for each database for the years 1990-2017 and also for each year from 1990 to 2017. These searches were conducted such that the search term of interest appeared at any point in the body of the text. However, this meant that the relevance of the sources or the extent to which a publication focused on the CEFR was not accounted for: the focus on the CEFR could range from a single mention of it at some point in the body of the work, or it could be a specific study about its usage or implementation. In the current study, these two examples contributed equivalently to the counts of articles on the CEFR, while they clearly make vastly different contributions to knowledge on the CEFR. As a result, a second search with the keyword in the title was also conducted, with the assumption that these publications focused more specifically on the CEFR. The first search intended to provide more comprehensive and inclusive results, while the second would provide results reflecting research with a deeper focus on the CEFR. The findings from both searches were considered in assessing the impact of the CEFR.

2.5 Source and geographical location

For the bibliometric indices of source and geographical location, a sort and filter tool on EBSCO Host was employed for the articles for which location metadata was available. This provided a list of journals and countries that contained or produced publications on the CEFR. Of the 12,104 hits that were retrieved on EBSCO Host, the metadata of 2,171 of them made up the results. For source, journal impact factor obtained from each of the journal's homepages, if available, was also noted (for a discussion about journal impact factor, see Garfield 2006).

2.6 Most cited works

Sort tools within the software Publish or Perish were used to rank the works with the greatest numbers of citations according to the retrievals on Google Scholar. Citations per year were also provided. The results of the two searches with the keyword in the body of the article or the title of the article are provided.

3 Results

3.1 Number of publications

A Google Scholar search of 'Common European Framework of Reference' for the years 1990-2017 retrieved approximately 18,400 publications. The EBSCO Host search for the same time period and search term produced a total of 12,104 hits. When the search criteria was restricted to containing the search term in the title alone, rather than anywhere in the article, EBSCO Host retrieved 305 articles, and Google Scholar, by way of the PoP software, retrieved 454. The results should be interpreted as representative of the data available through the tools EBSCO Host and Publish or Perish, and subject to their limitations.

3.2 Publications per year

Figure 1 shows the number of publications per year for the keyword 'Common European Framework of Reference' for the searches in each database. As can be seen in Figure 1, there are fewer than 10 publications in each year between 1990 and 1995. A gradual increase in publications between 1995 and 2001 is evident (from 10 in 1995 to 92 in 2001). In 2001, the number of publications jumps to 128. A gradual increase proceeds until 2013, with nearly 2,500 publications in that year. The number of publications increases slightly to over 2,500 in 2014 and 2015, peaks at nearly 3,410 in 2016, and then drops back to 2,810 in 2017. These patterns are similar in the literature searches in EBSCO Host until 2011. After 2011, the number of publications per year falls between 1,000 and 1,500 for each year thereafter and no increase per year in publications is visible (Figure 1).
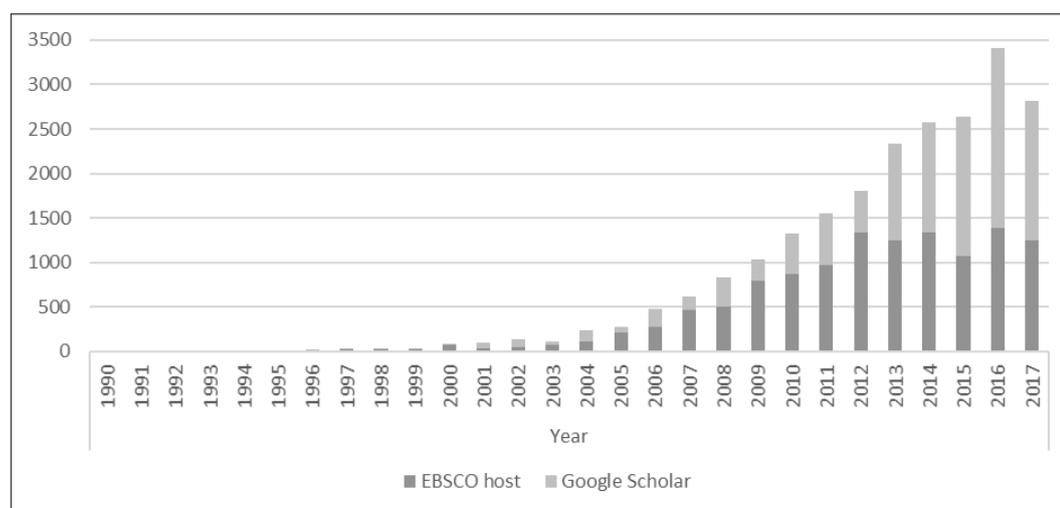
Figure 1. The number of publications for the search term 'Common European Framework of Reference' for each of the years from 1990 to 2017.

The results for the second search of works including CEFR in the title are shown in Figure 2. As can be seen, there are far fewer publications in each year when compared to Figure 1, although an increase of works over time, albeit a far less consistent one, is nonetheless evident. Once again, there are very few publications on the CEFR between its formal conception and the release of the first draft in 1995, with an increase in subsequent publications in the years until 2003. The number increases to over 20 works in the year 2004 and remains between 20 and 40 publications per year between 2004 and 2017, with the exception of the spike in 2012.
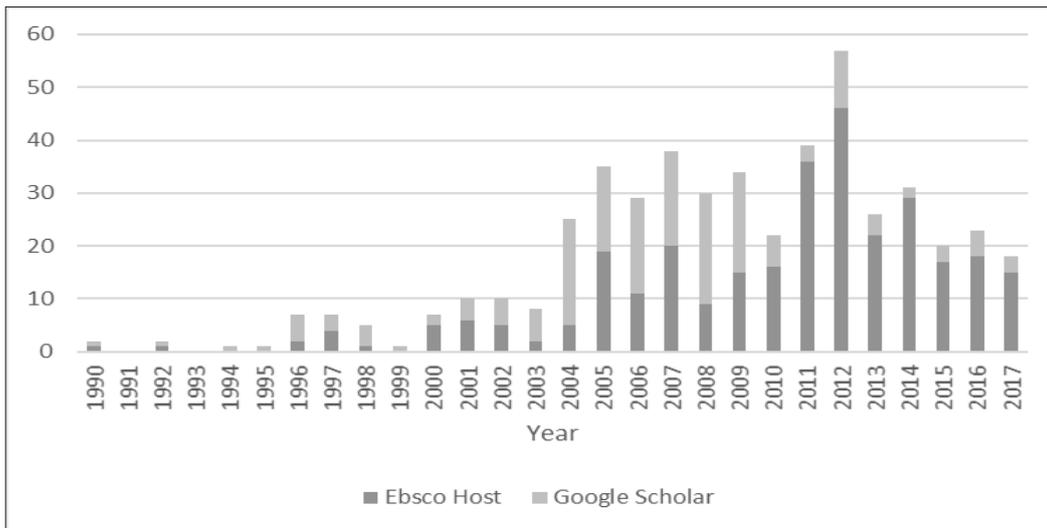
Figure 2. The number of publications for the search term 'Common European Framework of Reference' in the title for each of the years from 1990 to 2017.

3.3 Source and geographical location

The EBSCO Host search retrieved a total of 48 journals that published research on the CEFR ranging from 1 to 538 articles in each of these journals. The ten journals publishing a greater number of articles on the CEFR are shown in Table 1. Altogether, the top ten journals contained 1,714 relevant CEFR articles (nearly 80 percent of the total for which metadata were available). They are mostly published in English, with the exception of the 6th ranked journal, which contains mostly German language material.

EBSCO Host retrieved geographical information for 1,409 separate works. Three-quarters of these were European, including countries such as the U.K., Poland, Spain, Germany, France, Netherlands, Greece, Ireland, Italy, Finland, as the most common. Asia made up 11 percent of the remaining publications with the most research in Turkey, China, Japan, India, and Malaysia. Research from North America was mostly from the U.S. with about 30 percent from Canada. The countries of note from South and Central America and Oceania were Colombia and Australia respectively. In total, about 50 countries were identified where research on the CEFR was undertaken.

Table 1. *The ten journals with the highest number of articles on the CEFR according to an EBSCO Host search for the years 1990-2017*

Source	Number of articles	Impact factor (when available)
Modern Language Journal	538	1.745
Language Testing	228	1.815
ELT Journal	156	1.125
Language Assessment Quarterly	119	1.02
Language Teaching	105	1.913
Teaching German/Die Unterrichtspraxis	88	
Language Learning Journal	81	
Canadian Modern Language Review	77	0.39
Language Learning	68	2.079
European Journal of Language Policy	66	
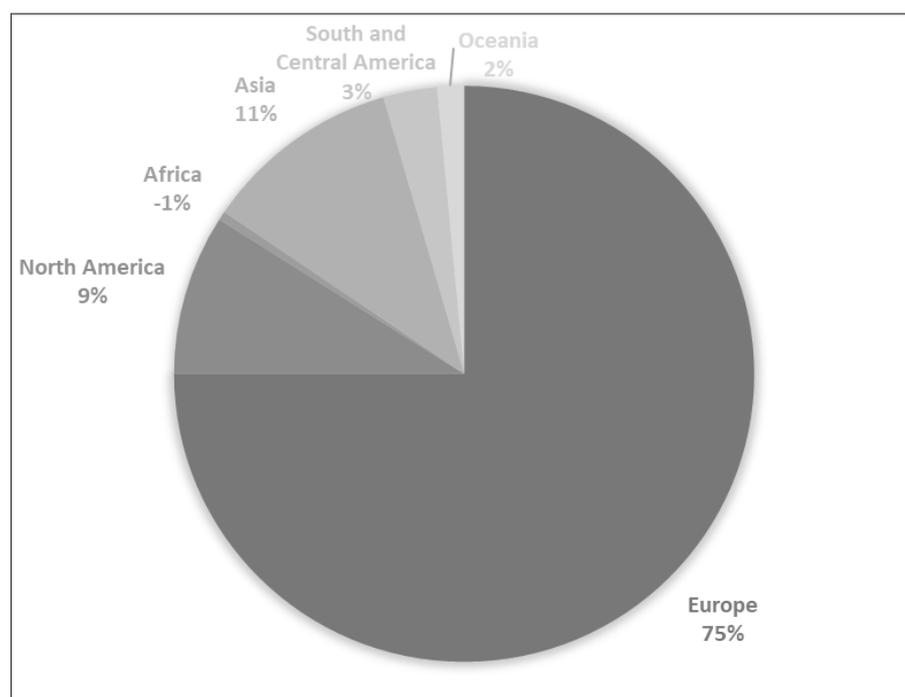
Figure 3. The geographical location of research on the CEFR according to the search term of 'Common European Framework of Reference' on EBSCO Host for the years 1990-2017.

3.4 Most cited works

Publish or Perish was used to identify the most cited works. The first 998 papers from the search with the search term appearing at any point were cited a total of 54,260 times. The 454 papers with CEFR in the title were cited a total of 3,029 times. The most cited ten publications with the CEFR at any point in the work are in Table 2, which also shows the number of citations per year since publication. Table 3 shows the most cited works with CEFR in the title alone. The framework itself is the only document to appear in both lists.

Table 2. *The ten most cited publications referring to the CEFR between 1990-2017*

Total cites	Cites per year	Authors/editors	Title	Year	Source type
6,664	952	C Baker	Foundations of bilingual education and bilingualism	2011	Book
4,176	2,088	V Cook	Second language learning and language teaching	2016	Book
1,731	432.75	J Jenkins, C Leung	English as a lingua franca	2014	Book
946	94.6	N Schmitt	Instructed second language vocabulary learning	2008	Article
885	55.31	M Byram, B Gribkova, H Starkey	Developing the intercultural dimension in language teaching	2002	Book
794	794	A Pym	Exploring translation theories	2017	Book
699		Council of Europe	Common European Framework of Reference for Languages: learning, teaching, assessment	2001	Document

Total cites	Cites per year	Authors/editors	Title	Year	Source type
566	35.38	D Marsh	CLIL/EMILE-The European dimension: Actions, trends and foresight potential	2002	Book
513	102.6	JE Purpura	Assessing grammar	2013	Book
487	97.4	M Byram, A Hu	Routledge encyclopedia of language teaching and learning	2013	Book

Table 3. *The ten most cited publications containing 'Common European Framework of Reference' in the title.*

Cites	Cites per year	Authors	Title	Year	Source
699		Council of Europe	Common European Framework of Reference for Languages: learning, teaching, assessment	2001	Document
185	16.82	D Little	The Common European Framework of Reference for Languages: Perspectives on the making of supranational language education policy	2007	Article
172	14.33	D Little	The Common European Framework of Reference for Languages: Content, purpose, origin, reception and impact	2006	Article
160	13.33	JC Alderson, N Figueras, H Kuijper, G Nold et al.	Analysing tests of reading and listening in relation to the Common European Framework of Reference: The experience of the Dutch CEFR Construct Project	2006	Report
121	5.76	JLM Trim	Modern languages: Learning, teaching, assessment: A common European framework of reference: A general guide for users: Draft 1	1997	Document
105	17.5	JA Hawkins, L Filipović	Criterial features in L2 English: Specifying the reference levels of the Common European Framework	2012	Book
84	14	M Byram, L Parmenter	The Common European Framework of Reference: The globalisation of language education policy	2012	Book
80	3.81	M Byram, G Zarate, G Neuner	Sociocultural competence in language learning and teaching: Studies towards a common European framework of reference for language learning	1997	Book
69	4.93	JC Alderson, N Figueras, H Kuijper, G Nold, S Takala	The development of specifications for item development and classification within The Common European Framework of Reference for Languages	2004	Report
61	8.71	Little D	The Common European Framework of Reference: A research agenda	2011	Article

4 Discussion

A bibliometric analysis was performed on research on the CEFR from 1990 to 2017, with the purpose of exploring the extent, provenance and adoption of the collected body of knowledge. In terms of the extent of the research, the results show a marked increase in the number of publications over the examined time, from 1990 to 2017 (Figure 1 and 2). The results suggest that there was scholarly interest in the CEFR following its formal inception in 1990, after the release of the first draft in 1995, and also in research conducted since the CEFR's publication in 2001. This means that greater attention is being paid to the CEFR from individual researchers and a greater number of researchers overall (Lockwood 2007). A peak in publications in 2016 was also seen, which may be due to the occurrence of Council of Europe language conferences held in October 2015 and March 2016 (Council of Europe 2015, 2016) and one specifically on the CEFR in Japan in March (FLP SIG 2016).

In addition to an increase in the overall number of publications, it was found that a range of journals publish work on the CEFR. These journals varied in their impact factor, geographical location, discipline, specific topics of focus, and even their main language of operation, thus suggesting that the CEFR has application in many areas within language education. When the geographical information of the publications was examined, the vast majority of the works (75%) were European, with research performed in North America and Asia making up nearly all of the remaining quarter. This suggests that the framework, while originally written for the European context, has utility in contexts outside of where it was developed.

In terms of the most cited works, the CEFR itself appeared at the top of the lists whereby the search term could appear either at any point in the publication or within the title of the work itself (Tables 1 and 2). For the former, as can be seen in Table 2, the most highly cited works were primarily books on a range of topics in language education and are not likely to focus greatly on the CEFR (which confirmed the rationale behind performing the second search with CEFR in the title). These findings suggest that scholars in language education are aware of and see value in the framework enough to discuss it or at least mention it in a wide range of works of varied topics. Conversely, for the works with CEFR in the title shown in Table 3, although the framework itself is the most cited work from this list, there is a wide range of source types (books, articles, and reports) and foci of the works: from language education policy, language testing, CEFR impact, and determining language proficiency (future studies could focus more closely on the thematic areas of research upon which the CEFR has been studied most extensively). This suggests that the CEFR has met its intended criteria, in the sense that its multi-purpose approach to language education is to be transparent, comprehensive, and cohesive (Council of Europe 2001). This also suggests that awareness of the CEFR is spreading, and that this has not only been occurring since it was originally published, but also more recently. This is also evident considering that the research from geographical locations external to Europe (and particularly Asia) is more recent than much of the European work. The works in Table 3, which contain the search term in the title, are also, on average, older than those presented in Table 2. This implies that the knowledge of the CEFR is increasing over time and that its uptake is occurring in contexts beyond where the CEFR was originally developed. In summary, the CEFR's impact appears to be spreading more and more widely as time goes by.

Although it has been shown that the amount of research on the CEFR has changed over the period of examined time, the characteristics of that change also have implications for the CEFR's impact. In Figures 1 and 2, a gradual and continual increase in publications from 2001 through to 2017 is mostly but not entirely evident. A tapering off of the growth in the number of publications can be seen in both figures, with local spikes at certain times. In Figure 1, the number of publications exceeded 2,500 in 2014, it did not increase significantly in 2015, went up in 2016, and then returned closer to 2,500 in 2017. In Figure 2, the number of publications remained between approximately 20 to 40 per year (with the exception of 2012) and dropped below this range after 2014. It is unclear whether the number of publications is in decline after 2016. If publications per year have declined or shortly will begin to decline, this could suggest that the framework has already had its greatest scholarly impact. However, this is unlikely given

recent developments such as the updated descriptors released in 2018 and their associated conferences (Council of Europe 2018, 2018b), as well as the release of this CEFR-specific journal. If publications per year continue to be produced at similar levels, this may mean that interest in the Framework has reached a level that will only change if impacted by exceptional events or activity in the literature or industry, as is suggested in the local spike of 2016. For example, the local increase in the number of publications in 2012 (Figure 2) may be a result of immediate increased awareness of the CEFR in Japan due in part to the development and release of the CEFR-Japan (Negishi et al. 2013). A national television station in Japan (Nihon Hoso Kyokai or NHK) adopted the CEFR as the basis for their foreign language education programming (Tono and Negishi 2012) which was followed by an outpouring of related works in Japan (see Runnels 2015; O'Dwyer et al. 2017). If the number of publications is still increasing, then the CEFR's full impact is yet to be seen.

In either case, each of these scenarios have implications for the extent of adoption of CEFR (Yeo et al. 2015), which may be better explored using a theoretical framework. Rogers' diffusion of innovation, a theory that seeks to explain the transfer of ideas, practices or items spread through communities and populations, offers such an opportunity for exploration. According to Rogers (2003), an innovation is communicated to members of social systems: whether the members adopt the innovation is dependent on the characteristics of the innovation and the individual. Specifically, members of the social system can be classified in five adopter categories, depending on their willingness to adopt the innovation, or their innovativeness. The adopter categories are often represented graphically on a bell-curve with time on the x-axis and market share on the y-axis (Rogers 2003) and have been found to make up consistent percentages of the social systems. The categories are innovators (2.5%), early adopters (13.5%), early majority (34%), late majority (34%), and laggards (16%). It should be noted that this refers to adopters only and not those that reject the innovation entirely, such that it does not include all members of a population. Furthermore, there is no assumption that once an innovation is adopted by a certain group it will continue to diffuse through the remaining categories; rather, diffusion can halt outright at any time.

The shapes of the curves of the bibliometric indicators (number of papers published by year, for example) can be used to explore the saturation and impact of an innovation within its industry, or to estimate its potential impact in the near future (Yeo et al. 2015). Furthermore, since changes in slope are associated with various levels of productivity (Koskinen et al. 2008), the results can be used to predict the degree and stage of an innovation's adoption. Indeed, the slope of the curve in Figure 1 changes in 1995, in 2001, and a third change is evident at approximately 2005 to 2006. These changes match up relatively well with CEFR-related events, namely the first draft's release in 1995 and the CEFR's release in 2001. During this period, the developers worked on the framework until the first draft in 1995, when it is possible that innovators began publishing research, followed by the contributions of early adopters between or shortly after publication in 2001 until about 2006. Indeed, this even matches up with the focus of a forum held in 2007 that was to go beyond the series of seminars and events introducing the CEFR and the potential it offers as a new approach to language learning, teaching and assessment (Goullier 2007), suggesting that it was intended for those who had already adopted the framework. The slope between 2007 and 2017 shown in Figure 1 can be interpreted in two ways: firstly, that there are two or three changes within that time, which suggests that the CEFR went from early majority from 2007 to 2012, to late majority in 2013, until it reached the laggards in 2016, and is in decline as of 2017, from having filled its market share (Rogers 2003). Realistically, the CEFR is very unlikely to have already reached laggard-adopters in any language education context in the world, and so the second and more likely possibility is that the slope can be seen as remaining consistent (with some local variations due to the influence from other geographical areas such as was discussed for Japan and the CEFR-J) from about 2012 onwards. This is supported by the EBSCO Host results, which also do not show much variation in numbers after 2012.

Some insight is gleaned when considering the results summarizing the number of works with CEFR in the title: Figure 2 shows a certain level of productivity from 1990 to 1995, another level between 1996

and 2003, and a third level after 2003, which arguably continues through until 2017. We know that the majority of these works are Europe-based, and due to CEFR being in the title we can assume that the research is performed by CEFR-adopters. These findings suggest that at least two, possibly three levels of adoption have occurred: the European innovators became involved after the publishing of the first draft and the early adopters started publishing two years after the CEFR's publication. It is possible that, currently, the early adopters are still the only ones publishing the same amount as when they first adopted the framework, but taking the findings from Figure 1 into consideration, it is more likely that diffusion into the early majority stage seems to have occurred and is ongoing at the time of writing.

Overall, this means that it took around or just over ten years after publication to move beyond the innovators and early adopters into the early majority stage in Europe, and following the normal-curve (Rogers 2003), this suggests it will take another ten to fifteen years for it to move beyond the late majority to the laggards (assuming no fundamental changes to the innovation or the social system). Although this accords with the timing cited in other innovation research works (Grübler 1996), in North America, for example, the CEFR is unlikely to have gone beyond the innovators. One reason for this is that the U.S. and Canada share an official language (compared to the numerous languages in Europe). They also have their own frameworks (ACTFL's Proficiency Guidelines in the US and the Canadian Language Benchmarks in Canada; American Council on the Teaching of Foreign Languages 2012, Citizenship and Immigration Canada 2013), which have been in operation since 1986 and 1996 respectively, and the need for the CEFR is lower (although arguments for its usage have been put forward in Canada, Arnott et al. 2017, Faez 2012: a Common Framework of Reference for Languages in Canada, a Canadian equivalent of the CEFR, is already in use in some parts of the country [Government of Saskatchewan 2013]). This may also be the case for Oceania. In Asia the socio-cultural situation may be more similar to Europe in that different languages are spoken in each country, significant resources are invested in language education, and no overarching framework is well-established. As such, the literature suggests that the CEFR is currently at an innovators stage for Asia overall (O'Dwyer et al. 2017) and may be entering the early adopters stage in Japan (Schmidt et al. 2017). Turkey also is one country where the CEFR may be moving beyond the innovators, based on the amount of nationally run programs that have supported its usage (Yalatay and Gurocak 2016; Sülü and Kır 2014). The CEFR's influence will be more notable over the next ten years in particular, possibly mirroring its European impact during the time after its 2001 release. What is clear from these analyses is that the CEFR has diffused and will continue to diffuse through different contexts at different rates.

This discussion is extrapolated from the findings of the bibliometric analysis performed on published research on the CEFR, and although findings suggest that scholars have had and will likely demonstrate continued interest in the Framework, we would like to highlight the caveat that there is a difference between teachers and researchers in its adoption. While many researchers are language teachers and vice versa, not all educators perform scholarly research, and not all researchers have taught. Although the CEFR is a language education innovation in which CEFR-adopter teachers perceive value, the patterns of uptake or adoption among teachers may be different and are difficult to determine. One possibility is that there is more research on the CEFR than there is actual usage, while another is that there is more widespread usage of the CEFR than the research shows, meaning that its impact is even larger than estimated. That being said, we think that the results of the bibliometric analysis are strong indicators that can be reasonably applied to represent adoption among educators as well as researchers. However, we must also note that these findings are unable to determine whether or not the impact that the CEFR has had on both scholarship and research is a positive one: the apparent interest in the CEFR shown in the results could be in part due to criticisms of the CEFR derived from its adoption and subsequent negative impact. Further studies could aim to assess the nature of its impact more precisely.

A methodological consideration with this bibliometric analysis is that the two databases generated overall total numbers that were divergent from each other. Although this did not present any major issues, as the findings from both of them were similar, future investigations of this kind should give

consideration to results of bibliometric analyses with different databases, as these often present varying perspectives, which then need to be interpreted individually, particularly in the social sciences (van Raan 2000). Nonetheless, these findings should be taken as preliminary since Google Scholar is not a fully manually curated database, nor did our searches include complete manual searches (as they do, for example, in systematic reviews and other types of literature reviews). Errors such as duplicates were found in the retrievals themselves (for instance, the most highly cited work in Table 2 had over 40 separate entries in Publish or Perish, meaning that its citation rates are most likely underestimated), and in the summations of retrievals: a global search on Google Scholar 1990-2017 retrieved different numbers than each of the searches for each year added together in Publish or Perish). While we selected EBSCO Host for its more detailed bibliometric information and metadata and to address such issues, this database also has some limitations including access to data: the articles and metadata available to EBSCO Host users are conditional to the specific members' library subscription. EBSCO Host identified approximately 12,000 CEFR-related articles (compared to Google Scholar's 18,000), and only a small percentage (about 20%) of the total articles and their metadata was accessible to the authors. It is possible that a different subscription could present different results. Despite these issues, the results likely provide a reasonable approximation of actual numbers, especially given that the patterning of results between the two databases were similar. We nonetheless warn that if the precise totals of publications are of importance, then other measures can be taken using alternative instruments and tools. We also suggest that future studies use different databases to perform searches, and modify and compare findings of different search terms and how research on the CEFR differs according to thematic area of study.

6 References

American Council on the Teaching of Foreign Languages. 2012. ACTFL Proficiency Guidelines. https://www.actfl.org/sites/default/files/pdfs/public/ACTFLProficiencyGuidelines2012_FINAL.pdf. (accessed 05 May, 2018.)

Alderson, J. Charles. 2007. The CEFR and the need for more research. *Modern Language Journal. 91*(4). 659-663.

Arnott, Stephanie, Lace Marie Brogden, Farahnaz Faez, Muriel Péguret, Enrica Piccardo, Katherine Rehner, Shelly K. Taylor & Meike Wernicke. 2017. The Common European Framework of Reference (CEFR) in Canada: A research agenda. *Canadian Journal of Applied Linguistics. 20*(1). 31–54. https://journals.lib.unb.ca/index.php/CJAL/article/view/24472/29666. (accessed 08 April, 2018).

Borgman, Christine L. & Jonathan Furner. 2002. Scholarly communication and bibliometrics. *Annual Rreview of Information Science and Technology. 36*(1). 1-53.

Carty, Nicola. 2014. The adult learner in Gaelic language-in-education policy. *European Journal of Language Policy. 6*(2). 195–217.

Citizenship and Immigration Canada. 2013. *National placement guidelines*. http://www.language.ca/wp-content/uploads/2017/12/National_Placement_Guidelines_2014-1.pdf. (accessed 08 April, 2018.)

Council of Europe. 2001. *The Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge, United Kingdom: Cambridge University Press.

Council of Europe. 2001b. *The common European framework of reference in its political and educational context*. http://assets.cambridge.org/052180/3136/sample/0521803136ws.pdf. (accessed 04 August, 2014.)

Council of Europe. 2005. *Survey on the use of the common European framework of reference for languages (CEFR): Synthesis of results*. Strasbourg, France: Council of Europe.

Council of Europe. 2018, May 16-18. Building inclusive societies through enriching plurilingual and pluricultural education. Strasbourg.

Council of Europe. 2018b. Companion volume with new descriptors. https://rm.coe.int/cefr-companion-volume-with-new-descriptors-2018/1680787989. (accessed 18 November, 2018.)

Council of Europe. 2016. 1ˢᵗ Symposium on *The linguistic integration of adult migrants: lessons from research,* Strasbourg. March 30-April. LIAM Project: www.coe.int/lang-migrants. (accessed 18 November, 2018.)

Council of Europe. 2015, October 14-15. Intergovernmental conference on *"The language dimension in all subjects: equity and quality in education"*. Strasbourg.

de Mejía, Anne-Marie. 2011. The national bilingual programme in Colombia: Imposition of opportunity? *Journal of Applied Language Studies. 5*(3), 7–17.

EBSCO Industries. 2018. *All databases*. https://www.ebsco.com/products/research-databases. (accessed 05 May, 2018.)

Eckes, Thomas, Melanie Ellis, Vita Kalnberzina, Karmen Pižorn, Claude Springer, Krisztina Szollás & Constance Tsagari. 2005. Progress and problems in reforming public language examinations in Europe: cameos from the Baltic States, Greece, Hungary, Poland, Slovenia, France and Germany. *Language Testing. 22*(3). 355-377.

Faez, Farahnaz, Suzanne Majhanovich, Shelley K. Taylor, Maureen Smith & Kelly Crowley. 2012. The power of "Can Do" statements: teachers' perceptions of CEFR-informed instruction in French as a second language classrooms in Ontario. *Canadian Journal of Applied Linguistics. 14*(2). 1–19.

Fagerberg, Jan & Bart Verspagen. 2009. Innovation studies: The emerging structure of a new scientific field. *Research Policy. 38*(2). 218–233.

Figueras, Neus. 2012. The impact of the CEFR. *ELT Journal. 66*(4). 477–485.

Framework & Language Portfolio SIG Newsletter. 2016. CriConCef III Symposium: Critical, constructive assessment of CEFR-based Language teaching in Japan and beyond. Osaka University, Nakanoshima Centre. March 26-27. https://sites.google.com/site/flpsig/critical-constructive-assessment-of-cefr/criconcefiii-march-2016. (accessed 05 May, 2018.)

Garfield, Eugene. 2006. The History Meaning of the Journal Impact Factor. *Journal of the American Medical Association. 295*(1). 90-93.

Glänzel, W., A. Schubert & H. J. Czerwon. 1999. A bibliometric analysis of international scientific cooperation of the European Union (1985–1995). *Scientometrics. 45*(2). 185–202.

Goullier, Francis. 2014. Le *Cadre européen commun de référence pour les langues*, instrument de normalisation ou document instrumentalisé pour une normalisation de l'enseignement et de l'évaluation? *Cahiers de l'Apliut. 26*(2).

Government of Saskatchewan. 2013. *A guide to using the common framework of reference (CFR) with learners of English as an additional language*. http://publications.gov.sk.ca/documents/11/82934-A%20Guide%20to%20Using%20the%20CFR%20with%20EAL%20Learners.pdf. (accessed 08 April, 2018.)

Hanney, Stephen R., Miguel A. Gonzalez-Block, Martin J. Buxton & Maurice Kogan. 2003. The utilisation of health research in policy-making: Concepts, examples and methods of assessment. *Health Research Policy System. 1.* 2–29.

Harzing Anne-Wil. 2007. Publish or perish, available from http://www.harzing.com/pop.htm. (accessed 12 April, 2016.)

Harzing, Anne-Wil & Satu Alakangas. 2016. Google Scholar, Scopus and the web of science: A longitudinal and cross-disciplinary comparison. *Scientometrics. 106*(2). 787–804.

Hulstijn, Jan H., J. Charles Alderson & Rob Schoonen. 2010. Developmental stages in second-language acquisition and levels of second-language proficiency: Are there links between them? *Communicative proficiency and linguistic development: intersections between SLA and language testing research.* 11-20.

Jones, Neil & Nick Saville. 2009. European language policy: Assessment, learning, and the CEFR.*Annual Review of Applied Linguistics. 29.* 51–63.

Karakaya, Emrah, Antonio Hidalgo & Cali Nuur. 2014. Diffusion of eco-innovations: A review. *Renewable and Sustainable Energy Reviews. 33.* 392–339.

Koskinen, Johanna, Matti Isohanni, Henna Paajala, Erika Jaaskelainen, Pentii Nieminen, Hannu Koponen, Pekka Tienari & Jouko Miettunen. 2008. How to use bibliometric methods in evaluation of scientific research? An example from Finnish schizophrenia research. *Nordic Journal of Psychiatry. 62*(2). 136–143.

Kostoff, Ronald N., Antonio del Rio, James A. Humenik, Ester Ofilia Garcia, & Ana Maria Ramirez. 2001. Citation mining: Integrating text mining and bibliometrics for research user profiling. *Journal of the American Society for Information Science and Technology. 52*(13). 1148–1156.

Li, Linda M. & George X. Zhang. 2004. A common framework for Chinese. Presented at Navigating the new landscape for languages (www.llas.ac.uk/navlang). 30 June - 1 July 2004. https://www.llas.ac.uk/resources/paper/2280. (accessed 13 April, 2016).

Little, David. 2006. The Common European Framework of Reference for Languages: Content, purpose, origin, reception and impact. *Language Teaching. 39*(3). 167-190.

Little, David. 2007. The common European framework of reference for languages: Perspectives on the making of supranational language education policy. *The Modern Language Journal. 91*(4). 645–655.

Lockwood, Fred. 2007. Foreword. In Gráinne Conole & Martin Oliver (eds.), *Contemporary perspectives in e-learning research: Themes, methods, and impacts on practice* (pp. xvi–xvii). London, United Kingdom: Routledge.

Martyniuk, Waldemar & José Noijons. 2007. Executive summary of results of a survey on the use of the CEFR at national level in the Council of Europe Member States. *Strasbourg: Council of Europe*.

Mays, Nicholas. Catherine Pope & Jenni Popay. 2005. Systematically reviewing qualitative and quantitative evidence to inform management and policy-making in the health field. *Journal of Health Services Research and Policy. 10*(Suppl. 1). S6–S20.

Milat, Andrew, Adrien Bauman, Sally Redman & Nada Curac. 2011. Public health research outputs from efficacy to dissemination: a bibliometric analysis. *BMC Public Health. 11.* 934.

Nagai, Noriko & Fergus O'Dwyer. 2011. The actual and potential impacts of the CEFR on language education in Japan. *Synergies Europe. 6.* 141–152.

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 Kraków conference. 135–163. Cambridge, United Kingdom: Cambridge University Press.

Ngo, Xuan Minh. 2017. Diffusion of the CEFR among Vietnamese teachers: a mixed methods investigation. *The Asian EFL Journal Quarterly. 19*(7). 9-32.

O'Dwyer, Fergus. 2014. Toward critical, constructive assessments of CEFR-based language teaching in Japan and beyond. *The Modern Language Journal. 91*(4). 659–663.

O'Dwyer, Fergus, Morten Hunke, Alexander Imig, Noriko Nagai, Naoyuki Naganuma, & Maria Gabriela Schmidt. 2017. *Critical, constructive assessment of CEFR-informed language teaching in Japan and beyond*. Cambridge: Cambridge University Press.

Papageorgiou, Spiros. 2014. Issues in aligning assessments with the common European framework of reference. *Language Value. 6.* 15–27.

Piccardo, Enrica. 2013. Assessment recollected in tranquility: The ECEP project and the key concepts of the CEFR. *Studies in Language Testing. 36.* 187–204.

Porter, Alan, Alisa Kongthon & Jye-Chyi Lu. 2002. Research profiling: Improving the literature review. *Scientometrics. 53*(3). 351–370.

Rogers, Everett. 2003. *Diffusion of innovations* (5th ed.). New York: Simon and Schuster.

Runnels, Judith. 2015. Usage of the CEFR and CEFR-J in Japanese universities: Preliminary survey results. *Framework & Language Portfolio SIG Newsletter. 14.* 8-18. https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnxmbHBzaWd8Z3g6MTBlODE1ZjJlINjg0NTYxZA. (accessed 12 April, 2016.)

Schärer, Rolf. 2007. The common European framework of reference for languages: Multifaceted and intriguing. *Babylonia. 1*(7). 11.

Sülü, Ayfer & Elif Kır. 2014. Language teachers' views on CEFR. *International Online Journal of Education and Teaching. 1*(5). 358–364. http://iojet.org/index.php/IOJET/article/view/69/97. (accessed 13 April, 2016.)

Runnels, Judith. 2014. The CEFR-J: The story so far (2012–2014). *Framework & Language Portfolio SIG Newsletter. 12.* 9-19. https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpb-nxm
bHBzaWd8Z3g6MTE5OTNmMWExOWQzNzdlOQ. (accessed 12 April, 2016.)

Schmidt, Maria Gabriela, Judith Runnels & Noriko Nagai. 2017. The past, present and future of the CEFR. In *Critical, constructive assessment of CEFR-informed language teaching in Japan and beyond*. Fergus O'Dwyer Morten Hunke, Alexander Imig, Noriko Nagai, Naoyuki Naganuma, and Maria Gabriela Schmidt (eds.). Cambridge: Cambridge University Press.

Tono, Yukio & Masashi Negishi. 2012. The CEFR-J: Adapting the CEFR for English language teaching in Japan. *Framework & Language Portfolio SIG Newsletter. 8.* 5–12. https://dl.dropboxusercontent.com/u/33808898/FLP%20SIG%20NL%208%20Sep2012%20CEFR-J.pdf. (accessed 11 April, 2016)

Tricco, Andrea C., Vivien Runnels, Margaret Sampson, & Louise Bouchard. 2008. Shifts in the use of population health, health promotion, and public health: a bibliometric analysis. *Canadian Journal of Public Health. 99*(6). 466–471.

University of Wisconsin–Milwaukee. 2014. *Diagnostic medical sonography: Journal impact factors & citation analysis*. http://guides.library.uwm.edu/c.php?g=56512&p=363435. (accessed 08 April, 2018)

Valax, Philippe. 2011. *The Common European Framework of Reference for Languages: A critical analysis of its impact on a sample of teachers and curricula within and beyond Europe.* Doctoral dissertation, University of Waikato.

Van Leeuwen, Thed. 2006. The application of bibliometric analyses in the evaluation of social science research. Who benefits from it, and why it is still feasible? *Scientometrics. 66*(1). 133-154.

Van Raan, Anthony F. J. 2000. The interdisciplinary nature of science: theoretical framework and bibliometric-empirical approach. *Practising interdisciplinarity.* 66-78.

Vine, R. 2006. Google scholar. *Journal of the Medical Library Association. 94*(1). 97–99.

Xian, Hanjun, & Krishna Madhavan. 2014. Anatomy of Scholarly Collaboration in Engineering Education: A Big Data Bibliometric Analysis. *Journal of Engineering Education. 103*(3). 486-514.

Yalatay, Sibel & Fatma Gurocak. 2016. Is the CEFR really over there? *Procedia. 232.* 705–712.

Yeo, Woondong, Seonho Kim, Hyunwoo Park & Jaewoo Kang. 2006. A bibliometric method for measuring the degree of technological innovation. *Technological Forecasting and Social Change. 95*, 152-162.

7 Biographies

Judith Runnels recently took leave from a graduate research degree at the University of Bedfordshire's Centre for Research in English Language Learning and Assessment (CRELLA). She currently works at an English language training center in France. She is interested in the usage of the CEFR, pluriculturalism, intercultural communication and learning oriented assessment.

Vivien Runnels PhD (Population Health) is a Senior Research Associate at the University of Ottawa who works in community-based research and evaluation, and globalization and health equity research. She has authored and edited publications for different audiences including municipal and provincial governments in Canada and community-based and international organizations.

How new CEFR mediation descriptors can help to assess the discussion skills of management students—Global and analytical scales

Irina Y. Pavlovskaya, St. Petersburg State University

Olga Y. Lankina, St. Petersburg State University

The article focuses on the assessment of mediation competence in the context of the Content and Language Integrated Learning (CLIL). We offer new assessment scales developed with the use of descriptors for mediation from the *CEFR Companion Volume* (2018). The approach to assessment of oral performance that we discuss combines global and analytical marks. For the majority of classroom teachers in Russia, this issue has become very important from two points of view: a) how to introduce new scales of mediation and connect them adequately with traditional speaking skills, described in the literature (Pavlovskaya 2017), and b) how to harmonize global assessment with analytical scales. The research is based on the experience of evaluating the mediation skills of students of the Graduate School of Management, St. Petersburg State University. The implications of the method for classroom teaching are discussed.

Keywords: mediation, oral performance, assessment, global and analytical marks, global achievement scale, analytical scale, CEFR descriptors, cognitive skills, relational skills, group discussion.

1 Introduction

CLIL teachers of management students always have to be on alert, looking for the 'soft' and 'hard' skills that students might need most. Mediation is partly a hard skill, because it is firmly based on proficiency in a foreign language as well as on the relevant professional knowledge, but it also covers the top 10 soft skills that are so attractive for employers (communication, flexibility, leadership, motivation, patience, persuasion, problem-solving abilities, teamwork, time management, work ethic) (*hard* skills *vs. soft* skills).

In our case, the aim of the classes is to develop language-related skills that managers may need at work. We think that facilitating and encouraging conceptual talks has become an important professional task of a manager. With this idea in mind, we focus on three task types: 1) how to facilitate discussions, 2) how to give persuasive talks, and 3) how to deliver business presentations. All of these tasks require mediation strategies.

Mediation, as it is defined in the CEFR Companion Volume (CEFR/CV), implies "passing on new information in an appropriate form; collaborating to construct new meaning; encouraging others to construct or understand new meaning, and creating the space and conditions for communicating and/or learning." (CEFR/CV 2018: 99). We also adopted the approach to learning as described by Brian North (North 2016: 9), who states that learners, and especially those who learn a foreign language, are usually confronted with the unknown, having to mediate new meanings to each other and thus find themselves challenged by situations that require reformulating a text or *mediating a text* (CEFR/CV 2018: 103-114). Alternatively, they have to *mediate concepts*, e.g. do problem solving, brainstorming and

concept development (CEFR/CV 2018: 114-119). The third type of mediation, *mediation of communication* (CEFR/CV 2018: 120-123), is less relevant to this particular environment, due to the fact that, linguistically and culturally, the students happen to be quite homogeneous.

It is crucially important to find an effective way of assessing the oral performance of students who are involved in group discussions on professional issues. To meet this challenge an empirical research setting has been employed using both global achievement and analytical scales.

2 Research setting

The research involves B.A. programme undergraduate students at the Graduate School of Management, St. Petersburg State University, Russia, and their teachers of English (See Table 1.)

Table 1. *Research participants*

1	Number of students	49
2	English language proficiency	B2+ / C1
3	Age	19 – 21
4	Department	Management
5	Teachers	3

The students speak advanced English and most of them have successfully passed IELTS, B2 First or C1 Advanced Cambridge exams. Within the university curriculum, they have two English classes a week, 90 minutes each. There are three teachers who have experience in rating speaking exams and who took part in a CEFR-linking project (familiarization, standardization training and cut-score setting). This background gives them a better understanding of new CEFR descriptors for mediation that are being used for assessment purposes within the research.

In the third and fourth semesters of their studies, students carried out a project on developing business plans for startups that they might launch in the future, for example, a family leisure club, online language courses, a waste collection company, a communal heating system or an urban park. Students worked in groups of three or four and presented their plans to the other groups. They facilitated discussions and gave persuasive talks. The most common classroom activity within this project was a discussion. During discussions, students informed their group members about the details of their business. For example, they explained how they created business budgets and estimated risks, or they asked for advice on how best to manage their startups.

The teachers tested the students at the beginning of the academic year to see how good the students were at holding group discussions. Then the students were divided into two cohorts, which we refer to as the 'Control Group' and the 'Experimental Group'. Both cohorts followed the standard program of English adopted by the University, but the Experimental Group did an additional component, which involved exercises in mediation and self-assessment with CEFR descriptors. Both cohorts had a similar time schedule of classes: four academic hours (45 minutes) per week, 15 weeks in a semester, which is a total of 120 hours per year. The discussions within the Experimental Group employed the techniques typical of mediating texts and mediating concepts, such as linking to previous knowledge, amplifying or streamlining the text, solving problems, inferring, etc. All of the students took an oral test at the end of the course.

3 Research question

The research question was as follows: How can we effectively integrate mediation into the set of criteria for oral assessment? We approached this question with the understanding that students complete a communication task successfully if they display good mediation skills. In addition, we expect them

to be intelligible, coherent and logical when presenting arguments, employ an appropriate range of grammatical patterns, have considerable lexical resources, and demonstrate sufficient accuracy of speech. Therefore, the analytical criteria should include (1) interaction, (2) discourse management, (3) range, (4) accuracy, and (5) phonological control. We also understand that the mediation, production and interaction skills are highly interdependent. Indeed, if students are not sufficiently intelligible or they have some problems with the accuracy or fluency of their speech, it would be highly unlikely that they could cope with a mediation task successfully.

The mediator reformulates, summarizes or streamlines information. At the same time he/she is trying to build rapport within the discussion group. That is why in order to assess mediation globally the assessor has to ask two questions: 1) has the student managed to convey information clearly, and 2) has the student facilitated the discussion and collaborated successfully to construct meaning? The answers to these questions help the assessor to decide on the global achievement mark for mediation. Consequently, the global achievement mark that evaluates the mediation skills describes (1) relaying information and (2) facilitating discussions and collaborating to construct meaning.

Keeping this in mind, we can suggest that the assessment of oral performance in a group discussion on professional issues would be effective if it includes awarding analytical and global marks, so that *five analytical marks* are given for 1) interaction, 2) discourse management, 3) range, 4) accuracy, and 5) phonological control, and the *global mark* is given for mediation.

4 Research methodology

The oral performance assessment scheme was developed for this purpose. Firstly, we outlined the skills of oral mediation that students need to acquire. In order to list the skills that we wanted to assess, we analyzed the needs of the students and mapped them onto the descriptors for mediation. We grouped *cognitive skills*, which cover relaying a text, shortening a text, and elaborating on the text (see Table 2), and *relational skills* (see Table 3), which refer to mediating concepts: facilitating collaborative interaction, collaborating to construct meaning, managing interaction, and encouraging conceptual talk (CEFR/CV 2018: 116-117; 119).

Table 2. *Cognitive skills*

1	*Relaying a text*
	• Can paraphrase and render its meaning.
	• Can adapt the style and change register to meet the needs of the recipient.
2	*Shortening a text*
	• Can highlight the key points.
	• Can choose the relevant information.
3	*Elaborating on the text*
	• Can link the issue to previous knowledge.
	• Can explain difficult notions.
	• Can explain relationships between ideas.
	• Can generalize to explain the meaning of examples.
	• Can provide examples to give meaning to abstract ideas.
	• Can use metaphors and idiomatic language to sum up.
	• Can transform complex notions used in the text into passages that are easy to understand.
	• Can speculate about the inferences used by the author.

Table 3. *Relational skills*

	Facilitating and managing collaborative interaction in groups
	Can define goals of the discussion.
	Can stimulate a discussion.
	Can steer a discussion towards a conclusion.
	Can conclude a discussion.
	Can show sensitivity to different perspectives in a group.
	Can organize a group discussion.
	Collaborating to construct meaning + encouraging conceptual talks
	Can present their ideas.
	Can invite reactions from other group members.
	Can further develop other people's ideas.
	Can participate in the group discussion accordingly, e.g. contributing to collaborative decision-making, highlighting issues, evaluating problems, elaborating points of view.
	Can encourage the other interlocutors to conduct conceptual talks.

Sets of tasks on professional topics were created for the training and final assessment. For the final assessment, students watched one of several videos on leadership; then they met in a group of five or six people who had watched different videos. They received a question for a discussion based on the problems raised in the video and the project that students were involved in. Students had to share their knowledge and experience about leadership styles, discuss a problem taking the role of a leader, and attempt to arrive at a conclusion. Those tasks were aligned to B2 CEFR level using the CEFR Grid for Speaking.

Finally, the criteria for the assessment scales were defined and their descriptors were adapted from those for B2 in the CEFR and CEFR Companion Volume. These descriptors were used in five-point analytical and global achievement scales for bands 1, 3 and 5.

Technically, the assessor listens to a group discussion (5-6 people), which continues for about 30 minutes and involves presenting the information that the students have researched or gained before. In addition, the students discuss conceptual issues. During the discussion, the assessor awards analytical marks to every student. After the discussion, the assessor gives students global achievement marks for mediation.

The discussions were recorded during the experiment. Subsequently, they were assessed by three raters. The first rater was the teacher, who conducted face-to-face assessments. The other two raters, also teachers, assessed the recorded performances. They used audio scripts to help identify students. These raters had undertaken tuning-in with standardized performances before assessing students' discussions. The aim of the tuning-in exercise is to remember what 'strong', 'average' and 'poor' performances are like and the raters did tuning-in exercises before each assessment session. The raters' correlation lay between 0.87 and 0.91 (Table 4).

Table 4. *Rater correlation*

RATERS	PEARSON
1 / 2	0.91
1 / 3	0.93
2 / 3	0.87

5 Results and discussion

The results of the experiment were statistically analyzed with the help of Excel and ITEMAN. The data provided by the three raters were collected and the average mark used further for calculations. The maximum score is 30.

Table 5. *Central trend measures and classical statistics for the two groups' scores*

	Statistics	Method of Calculation	Diagnostic Test		Final Test	
			Experimental Group	Control Group	Experimental Group	Control Group
1	Number of Participants	No program	24	24	24	24
2	Mean	Excel	20.63	17.35	21.11	17.54
3	Mode (bimodal distribution)	Excel	18.33	Mode 1 15.00; Mode 2 18.00	18.67	20.00
4	Median	Excel	19.17	17.67	21.17	17.50
5	Standard Deviation	Excel	4.03	1.88	3.40	2.40
6	Skew	Excel	0.32	-0.02	0.13	-0.07
7	Kurtosis	Excel	-1.02	-1.22	-0.95	-1.41
8	Min Score/ Max Score	No program	13.00/27.00	13.00/ 22.00	14.00/25.00	14.00/21.00
9	Mean Item	ITEMAN	No.	2.89	3.69	2.92
10	Alpha	ITEMAN	0.95	0.90	0.94	0.83
11	SEM	ITEMAN	1.00	1.02	0.97	1.01

High values of Cronbach's alpha, showing internal consistency of characteristics (Table 5, no. 10) and the measurement error not exceeding 1.02 (Table 5, no. 11) indicate the reliability of the test.

The values of the minimum and maximum scores (Table 5, no. 8) as well as the standard deviation (Table 5, no. 8) indicate a greater homogeneity of the Control Group in comparison with the Experimental Group. It should be noted that at the final test both groups demonstrated a more uniform level of skills development, which is confirmed by a decrease in the standard deviation.

Some heterogeneity in the population of the groups is indicated by the flat-topped distribution, expressed by a small negative Kurtosis (Table 5, no. 7).

The absolute value of the Asymmetry in both groups is not significant and does not exceed 0,32 (Table 5, no. 6), while remaining positive in the Experimental Group and negative in the Control Group. This may indicate the presence of several students in the Experimental Group who are demonstrating higher level of skills development and some students in the Control Group with a lower level. Nevertheless, we can state that the difference between Experimental and Control groups did not exceed 0.54 points at the beginning of the experiment and 0.77 points at the end (Table 5, no. 9), and is not significant for the purposes of our experiment.

Further statistical characteristics of the holistic criterion 'mediation' and analytical criteria 'interaction', 'discourse management', 'variability', 'correctness' and 'phonological control' were calculated with the help of the program ITEMAN.

For the Diagnostic and Final tests, the correlation of scores by the six criteria with the Mean score (Table 5, no. 2) were calculated (see Table 6).

Table 6. *Criteria scores and mean score correlation*

Criteria	Experimental Group. Diagnostic Test	Experimental Group. Final Test	Control Group. Diagnostic Test	Control Group. Final Test
Mediation	0.83	0.81	0.86	0.80
Interaction	0.91	0.87	0.84	0.72
Discourse Management	0.93	0.87	0.70	0.58
Range	0.88	0.85	0.72	0.71
Accuracy	0.88	0.79	0.87	0.64
Phonological Control	0.76	0.86	0.49	0.38

As we can see in Table 6, almost all criteria scores strongly correlate with the Mean, except for the phonological control, which is not surprising, as pronunciation does not necessarily correlate with overall communicative proficiency. It is noteworthy that the Experimental and Control groups differ in the way that mediation, interaction, and discourse management statistics changed from diagnostic to final tests (see Figure 1).

Figure 1. Diagnostic test (blue) vs. Final test (red) results in a) Experimental and b) Control groups
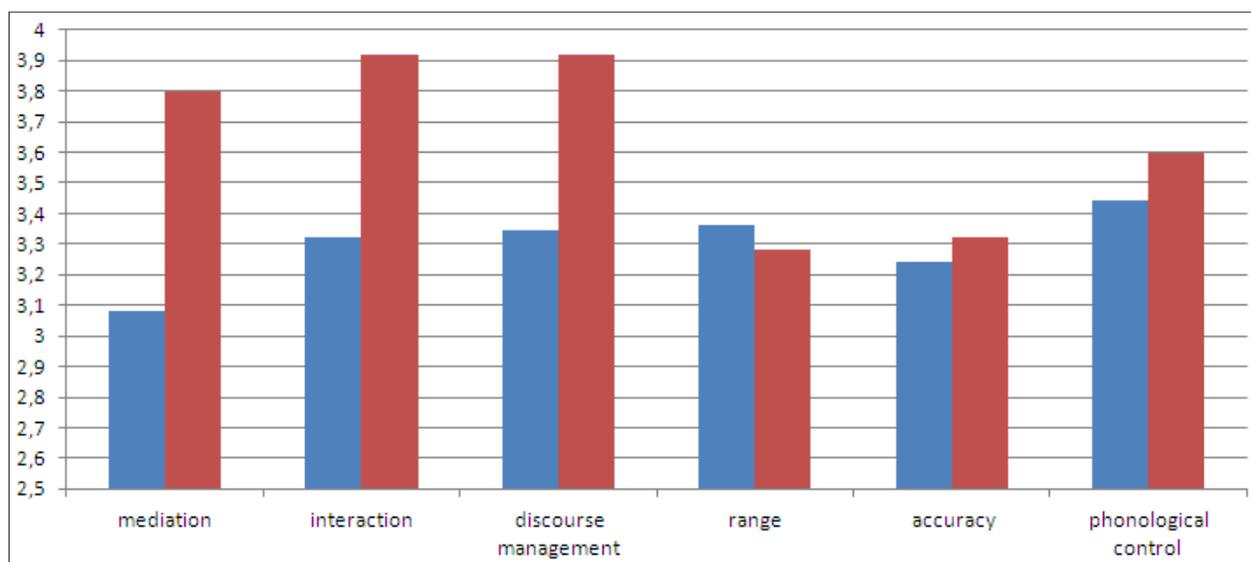
a) Experimental Group
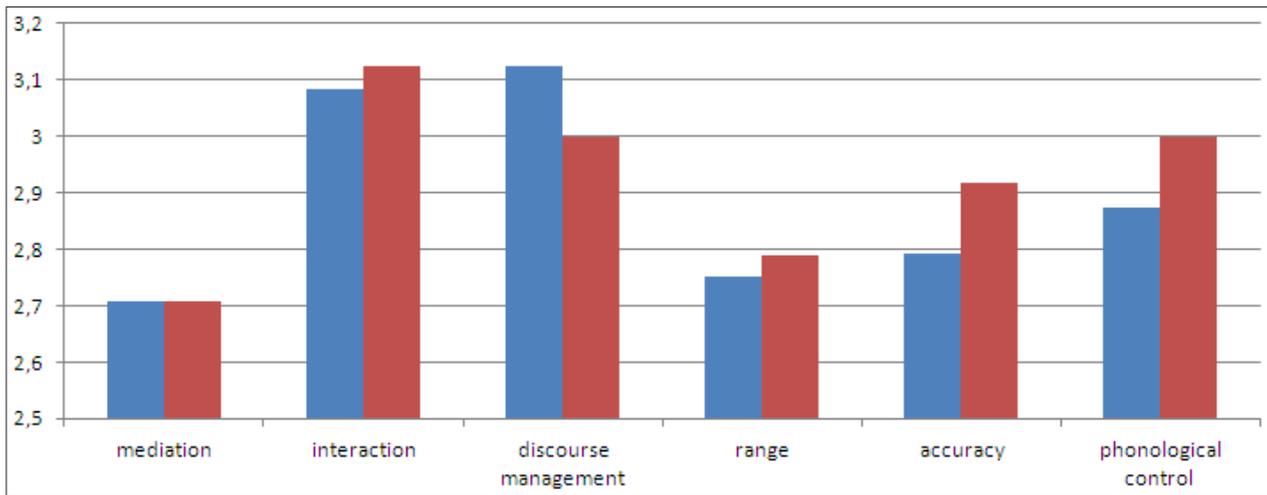
b) Control Group

Figure 1 presents the data of the diagnostic test (blue) versus the final test (red) in the two groups. The upper bar chart shows the results obtained from the group that had some additional practice with descriptors for mediation and the lower bar chart gives information about the group, which did not have this additional practice. From left to right we have twin bars of mediation, interaction, discourse management, range, accuracy, and phonological control. We can see that the performance of the group who worked with CEFR descriptors is slightly better than in the Control Group. This difference is quite small, but consistent. The difference is also stronger in relation to the three communicative criteria as opposed to the three linguistic criteria. These data may indicate the effectiveness of a set of exercises for the development of oral mediation skills in group discussion that was used in the Experimental Group.

6 Conclusion

The main conclusion is that the global achievement mark for mediation and the analytical marks are interrelated and we can support our analytical marks with the global mark for mediation and vice versa. To some extent, this approach can be regarded as *efficient* because it helps the assessor to self-check. The main implication of shifting from teaching communication to teaching mediation is the *increased focus on the collaborative development of new ideas*. By elaborating the concept of mediation and introducing mediation activities into the classroom, we facilitate passing on and receiving knowledge, and, most importantly, increase the autonomy of learners.

We realized that 'leading group work: encouraging conceptual talk' is a kind of activity that is often thought to be the responsibility of teachers, whereas the CEFR urges us to include it in students' repertoires, thus making them more independent. Working with CEFR descriptors can improve their social and collaborative skills.

Apart from these conclusions, some other interesting observations were made. For example, we noticed how mediation abilities develop with the progression of CEFR levels. At B2 level, students normally cannot grasp the totality of a complex abstract idea. Rather, they isolate two or three notions and explain them. At higher levels, students can mediate a concept in all its complexity as a whole. This could be a good indicator of students' level of language proficiency.

7 References

Association of Language Testers in Europe (ALTE). 2014. *The CEFR Grid for Speaking.* https://www.alte.org/resources/Documents/ALTE_CEFR_Speaking_grid_tests2014.pdf (accessed 6.11.2018).

Council of Europe. 2018. *The CEFR Companion Volume with new descriptors.* https://rm.coe.int/cefr-companion-volume-with-new-descriptors-2018/1680787989 (accessed 6.11.2018).

Council of Europe. 2001. *Common European Framework of Reference for Languages: Learning, teaching, assessment (CEFR)*, 2001. Strasbourg: Council of Europe. https://rm.coe.int/16802fc1bf (accessed 6.11.2018).

Hard Skills vs. Soft Skills: What's the Difference? https://www.thebalancecareers.com/hard-skills-vs-soft-skills-2063780 (accessed 6.11.2018).

North, Brian and Enrica Piccardo, 2016. *Developing illustrative descriptors of aspects of mediation for the Common European Framework of Reference (CEFR).* https://rm.coe.int/developing-illustrative-descriptors-of-aspects-of-mediation-for-the-co/1680713e2c (accessed 6.11.2018).

Pavlovskaya, Irina (ed.) 2017. *Актуальные вопросы языкового тестирования. [Modern issues of language testing].* Vol.2. St. Petersburg. http://hdl.handle.net/11701/7188 (accessed 08.04.2019).

8 Biographies

Prof. Irina Y. Pavlovskaya, Doctor of Philology works as a professor at St. Petersburg State University, Department of Foreign Languages and Linguodidactics. Her fields of research include language testing, language teaching methodology, phonetics and phonology, and phonosemantics. She has published more than 100 research and pedagogical articles. She supervised 13 successfully defended PhD papers (candidates' dissertations). She is a member of ILTA, EALTA, IATEFL (International Language Teaching Association), and the founding president of SPELTA (St. Petersburg English Language Teachers' Association).

Olga Y. Lankina of St. Petersburg State University holds a specialist degree in applied linguistics. She is a researcher, CLIL teacher and teacher of English for academic purposes at St. Petersburg State University, Russia. Olga is a Speaking examiner and examiner trainer. She is a member of EALTA. In 2013-2015 she was involved in the CEFR-linking project for the final test in English at St. Petersburg State University.

Implementing the CEFR at a Vietnamese university—General English language teachers' perceptions

Le Thi Thanh Hai, University of Foreign Languages, Hue University

Pham Thi Hong Nhung, University of Foreign Languages, Hue University

This paper reports preliminary findings regarding English language teachers' perceptions of the top-down implementation of the CEFR for non-English major students at a university in Vietnam. The study follows a mixed-method sequential design with the data being collected by means of questionnaire and interview. The findings have shown that General English (GE) language teachers have a sound understanding of the CEFR's values, think positively about its readiness and have relatively good awareness around the necessity for its implementation. Yet they express major concerns about the work and tasks involved in the CEFR implementation process. The most frequently cited reasons are associated with time constraints, limited access to relevant teaching materials and the tremendous gap between students' admission levels of proficiency and the expected CEFR-based learning outcomes. Relevant suggestions are drawn out with the hope of improving the process of implementing the CEFR in a specific context and facilitating fruitful educational changes to take place.

Keywords: perceptions, language policy, educational reform, EFL

1 Introduction

Soon after its publication in 2001, the Common European Framework of Reference (henceforth the CEFR) gained attention and respect, not only in Europe but also in the rest of the world (Alderson 2002, Byrnes 2007, Hulstijn 2007, Tono and Negishi 2012). The enthusiasm for the document has been recognized to extend far beyond Europe to Latin America, the Middle East, Australia and parts of Asia (Byram and Parmenter 2012). Outside the European context, as a "supranational language education policy" (Little 2007: 645), the CEFR has been observed to have major influences in language policy planning (Bonnet 2007, Byrnes 2007, Little 2007, Nguyen and Hamid 2015, Pham 2012, Pham 2017) especially in countries where English is taught as a foreign language. A number of Asian countries have experienced the implementation of the CEFR in national contexts as an attempt to reform the system of language teaching in the country. Vietnam is not an exception.

In 2008, the Vietnamese Government launched a national project named "*Teaching and learning foreign languages in the national educational system for the 2008-2020 period*", often referred to as Vietnam's National Foreign Languages 2020 Project (henceforth 2020 Project) as a national strategy so as to renovate the foreign language teaching and learning in the national education system during the period 2008-2020 (MOET 2008), now extended to 2025 (Vietnamese government 2017). The most significant part of the 2020 Project is the adoption of the CEFR, a global framework, into the local Vietnamese context of language teaching and learning as a "quick-fix" (Steiner-Khamsi 2004) solution to restructure the national foreign language education system. On the basis of the CEFR, a Vietnamese version of CEFR

was developed, approved and legitimated by Vietnamese authorities (MOET 2014). This CEFR-aligned framework is actually the translation of the CEFR into Vietnamese with very few modifications (Pham 2017, Pham 2018). The CEFR-based levels of proficiency were used to set standards for learning outcomes at different levels of education, from primary through secondary and high schools to universities. Students leaving primary schools at grade 5 are expected to achieve the CEFR-A1 level, lower secondary and high schools the CEFR-A2 and B1 respectively. Students majoring in English must achieve level C1 to be entitled to be granted university graduation degree while non-English majors must obtain B1 level.

The CEFR global levels were also utilized to set standards for teacher professionalism. Teachers teaching English at primary and lower secondary schools are asked to achieve B2. Those teaching English at high school or higher should obtain C1 and above. This adoption of the CEFR as the standard for both student outcomes and for professional assessment, underpinned by the 2020 Project in Vietnam, had been hoped to bring about positive, radical changes as is clearly stated in Decision 1400 of the government.

However, there have been warnings that the success of this ambitious language policy could be threatened by both its unfamiliar and top-down nature.

Firstly, since adapted from the CEFR, whose original purpose is not directed to diverse language contexts around the world but revolves around Europe, this alien framework may give rise to paradoxes if it is not carefully contextualized (Pham 2017). With remarkable differences in terms of social needs, language learning and teaching conditions, qualifications of language teachers and proficiency levels of learners as well as expectations and purposes, the appropriateness of the CEFR-aligned framework in Vietnam may be questioned. The implementation of the CEFR in Vietnam could, thus, be very socio-political in nature if "using the European model regardless of how inappropriate such a model might have been" (Kaplan and Baldauf 1997: 153).

Secondly, the Vietnamese CEFR-aligned framework has been forwarded to lower levels for implementation without explanatory reasons being given for its adoption (Pham 2017) nor with any consultation with the ultimate language learners and users. There is also a lack of previous research and pilot use of this framework in Vietnam (Pham 2012, Pham 2018). Even now, there is no official document or research evidence describing the involvement of teachers and students in the process of making decisions around applying the CEFR in Vietnam. When teachers' perceptions or their students' need and wants are not taken account, it is synonymous that teachers' ownership of innovation was denied and the possibility of teacher feedback was minimal (Kennedy 2013). As such, the adoption of the CEFR can be considered to follow the 'top-down' approach, clearly reflected in the literature on language planning. Accordingly, practitioners, especially teachers and learners at the lowest levels have had no say in this policy-making. Teachers are envisioned only as implementers of the policy and not as players of key roles in the centralized language planning processes (Poon 2000, Waters 2009). Therefore, the implementation of the CEFR in Vietnam is likely to create some mismatches between the expectations of adopters, those who sanction (government officials) the innovation and those who implement (teachers) it. The need for research, on the topic of the national adoption of CEFR language policy and issues of its implementation, has emerged.

In addition, research has shown that problems and failures in the implementation phase may emerge from teachers themselves due to their attitudes and behaviour. Although teachers' perceptions and attitudes are not always reflected in what actually teachers do in the classroom, they do influence practices (Borg 2009) and teachers' practices are considered as the visible part of the teaching iceberg (Waters 2009). In understanding teachers' perceptions, the submerged part of the iceberg can be of great importance in explaining what teachers do in the classroom. As for the implementation process, teachers, as implementers, play a significant role in bonding learners, materials, teaching practice and assessment altogether. However, studies have demonstrated that teachers do not always do as directed nor did they always act to maximize policy objectives (Cohen and Ball 1990, McLaughlin 1987). Additionally, they have been diagnosed as "resistant to change" (Wang 2008) or unwilling to implement a teaching innovation despite expressing positive attitudes towards it (Kennedy 1999, Keranen 2008,

cited in Waters 2009). Resistance, subversion and/or indifference are among the teachers' attitudes towards change and innovations.

Surrounding the implementation of the CEFR in Vietnam, the need to understand teachers' perceptions of, and responses to, this language policy implementation are obvious. Yet limited research has been conducted around this issue. The impacts on the language education system, on teachers' and learners' attitudes and perceptions toward the use of the CEFR, on the effectiveness of such changes in (foreign) language policy, have not been considered. As the implementation process is both comprehensive and profound, the need for more research on adopting the CEFR to Vietnam is clear. For that reason, this research is an effort to explore the CEFR implementation in Vietnam from the grass roots perspective.

2 The study

2.1 Research setting

The present study examines GE teachers' perceptions of implementing the CEFR at a Vietnamese tertiary setting as opportunities for understanding teachers' voices to a 'top-down language reform policy' (Nguyen and Hamid 2015, Pham 2017) in Vietnam. Given that the large-scale CEFR implementation applies to both English major and non-major curricula, this study chose to focus more on the CEFR-aligned General English curriculum for non-English major students and the challenges GE teachers face during the process of implementing this curriculum.

Hue University, where this research was conducted, is a regional university in Central Vietnam. Its non-English major students come from the Central Highlands and the provinces and cities in the centre of the country. According to their major field of study, students attend different colleges of Hue University with Hue University for Foreign Languages having full responsibility for English teaching to students from all colleges. Students vary in terms of social backgrounds, major fields of study chosen, and English proficiency, but most enter university at the age of 18 years. Teachers also differ in origin, experiences, qualifications and expertise. The Ministry of Education and Training (MOET) mandated that, as a state-run university, Hue University must have its non-English major students achieve CEFR B1 level as one condition for being granted a university graduation degree.

MOET stipulated Level 3 (equivalent to CEFR-B1 level) as the minimum language proficiency requirement for university graduation of non-English major students. Since MOET sets the learning outcomes for learners independent of curricula and teaching materials, the burden on the shoulders of state-run universities, teachers and students is heavy. MOET also compels a 7-credit general English curriculum be provided for non-English major students before their B1 CEFR-aligned examination. In effect, non-English major students have a total of 105 teacher-led hours of English classes in their first three semesters, divided into 30-30-45 hours respectively, and are expected to achieve level B1. In theory, the majority of those students have already spent seven to ten years learning English at school, so the expected B1 CEFR-aligned learning outcome should be achievable. The reality is different: large numbers of students leave high school without being able to speak any English at all although they may have accumulated relatively good knowledge of its grammar and vocabulary (MOET 2014b). It is therefore, not surprising that the non-English major students of Hue University vary greatly in their English proficiency levels.

2.2 Research question

The research aims to address the following question: What are GE teachers' perceptions of the CEFR and of its implementation for non-English major students?

Specifically, the study explores GE teachers' understandings of the values of the CEFR, their perceptions of the need for the CEFR implementation and its readiness for application in their context, and their perceptions of the work involved in the implementation process.

2.3 Participants

The study's focus on GE teachers' perceptions of implementing the CEFR for non-English major students at Hue University determines the inclusion criteria for participation. Forty-five (45) teachers who have experience in teaching GE for non-English major students for at least a semester were invited to participate in the study. Thirty-six (36) of these participated in the survey, giving a response rate of 80%. The remaining nine (9) teachers either refused or were absent on the day of questionnaire delivery. Eight (8) of the thirty-six (36) participants took part in the semi-structured interviews. Teacher demographic information is shown in Table 1, below.

Table 1. *Demographic data of participants*

		Count
Gender	female	29
	male	7
Years of teaching non-English major students	< 5 yrs	7
	6-10 yrs	4
	11-20 yrs	17
	> 20 yrs	8
Highest qualification	Bachelor	5
	Master	30
	Doctor	1
Another Bachelor degree in languages	No	24
	Yes	12
CEFR training attended	By MOET	11
	By home university	26

Note. The total number of participants was 36.

Of these thirty-six (36) teachers, twenty-four (24) confirmed that the information and knowledge they have about CEFR and its application policy came from workshops provided by their home university, eighteen (18) from self-exploration including learning from colleagues and eleven (11) had the opportunity to attend CEFR training workshops conducted by the MOET. This suggests that a number of participants have attended more than one workshop on the CEFR and its implementation.

2.4 Research instruments

2.4.1 The questionnaire

A questionnaire (see Appendix for full form of the questionnaire) was used to gain quantitative data on teachers' perceptions of the CEFR implementation. It was developed and modified from a pilot questionnaire. Except for the first five questions about teacher demographics, the other twenty-seven (27) questionnaire items are in closed format.

Specifically, the first part of the questionnaire consists of five (5) questions investigating teachers' gender, teaching experiences and qualifications. The remainder of the questionnaire contains 27 five-point Likert scale items eliciting teacher perceptions of the CEFR implementation for non-English major students at Hue University. All of the items are developed and designed on the basis of a careful literature review of the CEFR and its implementation in different contexts. The 27 items were further divided into four main clusters focusing on the participants' perceptions of *the values of the CEFR, the readiness for the*

CEFR application, the necessity of applying the CEFR and *the work involved in the CEFR application process*. The five-point scale is coded in accordance with the logical way of thinking that the bigger the number, the higher the level of agreement is; i.e. 5 stands for "*strongly agree*", 4 for "*agree*", 3 for "*no idea*", 2 for "*disagree*" and 1 for "*strongly disagree*". Participants were asked to tick the number representing their level of agreement. A summary of the questionnaire is provided in Table 2 below.

Table 2. *Summary of the questionnaire*

Teachers' perceptions of the CEFR implementation	Items
Values of CEFR	3, 5, 8, 12, 13, 18
Necessity of CEFR application	20a, 20b, 20c, 20d, 20e, 20f, 20g, 20h
CEFR readiness for application	2, 9, 11, 15, 17,
Work involved in CEFR application process	1, 4, 6, 7, 10, 14, 16, 19

2.4.2 In-depth interview

Interviews were employed to provide richer data to complement the closed format of the questionnaire and to focus more on exploring the reasons underlying the participants' perceptions. Interview data helped to provide more insightful information and deeper clarification into the reasons for teachers' choices, why they perceived things in certain ways and what contextual factors influenced their perceptions (Creswell 1998). Identified issues developed from the quantitative questionnaire data became the basis for more in-depth exploration. Each interview had two parts (see Appendix for main interview questions). The first part consisted of a preamble and demographic questions. The main aim is to provide the participants with general information related to the purpose of the study, explain the ethical issues and establish good rapport between the interviewee and the researcher as well as to gather some demographic information from the interviewee. The main part of the interview explores further teachers' perceptions of the CEFR implementation in their context. Ten (10) main questions were developed in line with the four (4) afore-mentioned clusters from the questionnaire. For each question, the researchers also prepared in case there was a need to elaborate more on the participant's ideas and reflections. The order of the questions could also change, dependent on the flow of the interview but the same interview protocol was used to serve as a reminder for the researcher about the procedure and purpose of the interview (Creswell 2013) and to ensure consistency between all participants. The data provided an insightful exploration of general English teachers' perceptions; why they perceived the CEFR implementation process that way and what factors may have affected their cognition and understanding.

2.5 Data collection process

The data collection procedure of the present study followed Creswell and Clark's (2007) mixed method sequential model. The procedure lasted nine months from April to December 2017, beginning with the survey questionnaire and in-depth interviews for the pilot phase in April and May. After two months spent analyzing the pilot data and revising the instruments, the official questionnaire and interview questions were ready by the end of August 2017 and the survey was conducted between September and December 2017.

After the questionnaire had been collected and analyzed, eight interviews were conducted with eight participants who had agreed to do so. Each interview lasted about thirty minutes. All the interviews were conducted in Vietnamese and recorded for later transcription. The interviews were then transcribed, coded, and analyzed. Two or three weeks after the interviews, the researcher sent the transcripts to each participant for checking. No participant requested any changes to the transcripts.

2.6 Data analysis methods

Data analysis was conducted carefully and with consideration to ensure the reliability and validity of the study. Quantitative questionnaire and qualitative interview data were analyzed separately using different techniques. Quantitative data from the questionnaire were dealt with first, using descriptive and analytic statistics, followed by qualitative findings from the transcribed interviews, coded into and counted by themes.

   After data from the survey questionnaire had been collected and raw data input had been carried out, data cleaning and data filters were applied to ensure the validity and reliability of the questionnaire. Cronbach Alpha value of .844 for the questionnaire was gained, proving the reliability of the questionnaire and data collected. To gather qualitative data from the interviews, these were transcribed and sent to the interviewees for accuracy checking, then the interview recordings were listened to many times and transcribed notes were read and reread, assisting in assuring the accuracy of the language captured by the transcribed notes. Simultaneously, participants' voices and tones were captured to gain deeper understanding of their perceptions and attitudes to the issues under investigation. As themes and sub-themes emerged from data analysis, a full list of corresponding themes was created. By doing this, researchers can find answers to the research questions and simultaneously develop a deeper understanding of the phenomenon (Creswell 2013). Qualitative findings from the interviews were used to triangulate with quantitative findings from the questionnaire and to verify quantitative findings against qualitative ones.

3 Findings and discussion

Firstly, the results of questionnaire data analysis are presented in Table 3 below.

Table 3. *General English teachers' perceptions of the CEFR and its implementation*

No	Items	Contents	N	Mean	Std. Deviation
The values of the CEFR			**36**	**3.97**	**.495**
1	3	CEFR can make learning outcomes transparent	36	4.19	.920
2	5	CEFR helps create mutual recognition across institutions	36	3.86	.798
3	8	CEFR encourages self-directed learning	36	3.92	.649
4	12	CEFR helps renew assessment practice	36	3.83	.878
5	13	CEFR can help renew curriculum	36	4.03	.774
6	18	CEFR can create positive changes in English language education	36	4.00	.632
The reasons and necessity of the CEFR implementation in Hue University			**36**	**3.60**	**.452**
7	20a	CEFR is a global comprehensive framework	36	3.94	.826
8	20b	The teachers involved in the process are ready	36	3.44	.843
9	20c	The students involved are ready	36	3.28	.914
10	20d	CEFR has been well applied in other countries	36	3.33	.676
11	20e	The university has all resources required	36	3.56	.877
12	20f	CEFR can help improve the teaching quality of the university	36	3.89	.708
13	20g	The university can promote its reputation	36	3.69	.822
14	20h	CEFR implementation will improve the language proficiency of the students of the university	36	3.69	.822

No	Items	Contents	N	Mean	Std. Deviation
The CEFR readiness for application			**36**	**3.71**	**.594**
15	2	The CEFR descriptors of proficiency levels are representative	36	4.06	.715
16	9	CEFR is English-specific	36	3.39	.934
17	11	CEFR is context- specific	36	3.33	.926
18	15	CEFR is ready for any curriculum renewal	36	3.61	.934
19	17	CEFR descriptors need to be specified	36	4.17	.878
The work involved in the CEFR application process			**36**	**3.19**	**.570**
20	1	Necessary resources for the implementation were provided	36	3.86	.833
21	4	The implementation of the CEFR was piloted	36	2.69	1.142
22	6	Capacity building for the implementation (e.g. training workshops on the CEFR) was provided	36	3.81	.889
23	7	Staff involved were informed about the CEFR values and limitations	36	3.67	1.095
24	10	All teachers were involved in the CEFR-aligned curriculum design	36	1.56	.558
25	14	Staff involved were trained for the application/implementation procedure	36	3.39	.964
26	16	Expertise and professional support during the implementation process were provided	36	2.56	1.027
27	19	The objectives were realistic within the required timeline	36	3.06	1.068

3.1 General results

The average mean values of the four clusters ranged from 3.19 to 3.97, between levels 3 and 4 of the five-point Likert scale, which indicated that GE teachers had neutral to relatively positive perceptions of the CEFR and its implementation for their non-English major university students. Specifically, the level of teachers' agreement regarding the CEFR's value reached close to 4.0 (M= 3.97) and were slightly higher than those given to the need for the CEFR's application and its readiness for implementation (M=3.60 and 3.71 respectively). Nevertheless, they perceived the work involved in implementing the CEFR process as the lowest with a mean value of only 3.19. Of note is the fact that the first three clusters related more to the CEFR itself while the fourth concerns its application to General English for non-English major university students. It can be concluded that GE teachers have a generally sound understanding of the CEFR and its use. However, their perceptions of the CEFR implementation process were not as high. The next sections will present detailed discussion of these clusters together with the themes and sub-themes that emerged from interviews.

3.1.1 GE teachers' understanding of the values of the CEFR

Details of teachers' perceptions of the values of the CEFR can be seen in Table 3 above. Specifically, their agreement that CEFR can make learning outcomes transparent, can renew the curriculum and create positive changes in English language education reached above 4 of the five-point scale (4.19, 4.03 and 4.0 respectively). Other purposes such as encouraging self-directed learning, creating mutual recognition across institutions and renewing assessment practice received the mean values below 4 on the five-point scale of agreement (3.91, 3.86 and 3.83 respectively). Attention is drawn to the mean values of items being quite close to the mean value for the whole cluster of 3.97, suggesting that GE

teachers well understood the comprehensive objectives and principles of the framework including their application to non-English major students.

Data from the interviews generally aligned with quantitative findings. Of eight respondents, six teachers claimed that the CEFR's overall objectives met Vietnam's need for integration in the current situation. They also supported MOET's aims that the language proficiency of Vietnamese could be improved through implementation of the CEFR. From their comments, GE teachers' understanding of the values of CEFR could be captured. In brief, they understood that the policy for non-English major students was part of the bigger picture of efforts to boost foreign language education nationwide, at different levels of education and in different contexts, not just within their university. One teacher emphasized the potential to create mutual recognition between institutions with the CEFR-aligned outcomes, which was a favorable condition for students pursuing education at another university or institution.

In their context of teaching General English to non-English major university students, four out of the eight interviewed teachers expressed satisfaction with the CEFR division of language proficiency into six skill levels with concise descriptors for each level and for different language skills. They believed that this made the learning outcomes more specific and transparent. One participant also added that the descriptors "aided teachers and students a lot as they could see more clearly what and how they should do to get through to the end of their teaching and learning journey by looking at the B1 CEFR-aligned learning outcome". In other words, the interviewed teachers believed that their English teaching and learning became better oriented through the CEFR implementation. This finding was in line with that of Pham (2017). Data from the interview sessions also showed that teachers were aware of the interdependence among different domains of language education from outcomes, assessment to teaching materials and pedagogy. This idea reflected one feature, previously pinpointed by Little (2006), of the CEFR's contribution to language education worldwide.

In sum, GE teachers had a sound understanding of the CEFR's values. This finding was similar to that of Pham (2017) but differed from that of Nguyen and Hamid (2016). In Nguyen and Hamid (2016), the value of the CEFR to teachers was limited to "testing scores and numbers only" (p. 69). This difference could be partly explained by the different timing of research, with theirs being conducted during the first years of the CEFR implementation program while the present study was carried out six years after its first implementation. Another explanation may arise from the difference between the participant groups, the former investigating English language teachers of both English major and non-major students while the latter focused on GE teachers of non-major students only.

3.1.2. GE teachers' attitudes towards the necessity of the CEFR implementation

On average, the mean value of the whole cluster fell between 3 (*no idea*) and 4 (*agree*) (M=3.60). Synonymously, GE teachers were aware that implementing the CEFR at their home university was required, although their level of agreement was not high. Specifically, they agreed that the application of the CEFR was necessary because it provided a comprehensive global framework (M=3.94) and applying the CEFR would help to improve teaching quality (M=3.89), promote the university's reputation (M=3.69) and improve students' language proficiency (M=3.69). But they did not fully agree that the teachers, students and the university's resources were ready for this implementation. The mean values were close to middle value of 3.0 for the readiness of students, teachers and the university resources (M=3.28, M=3.44, and M=3.56 respectively) and indicated that teachers did not agree that their university was ready for such an application. In addition, they did not support the idea that it was necessary to apply the CEFR in Hue University because the framework has been successfully applied in other contexts (M=3.33).

There are two issues worth noticing from the quantitative results regarding GE teachers' perceptions of the necessity of the CEFR implementation. Firstly, all items showed high standard deviations (SD), with values ranging from .708 to .914, showing an ambit of teachers' viewpoints. In other words, GE

teachers' perceptions differed widely. Although the mean values of some items are quite high, it cannot be concluded that every teacher shared the same level of agreement. Secondly, the mean values varied greatly among items, revealing that the teachers had different perceptions regarding the necessity of applying the CEFR to non-English major students at their university.

Items related to the potential impacts and effects of the CEFR implementation, such as on the school's reputation, promotion, teaching quality and students' proficiency improvement received relatively positive rankings. In comparison, the items concerning school infrastructure and capacity readiness obtained a much lower level of agreement from GE teachers.

From the findings, it can be concluded that GE teachers' positive perceptions of the necessity of implementing the CEFR came mainly from their trust in the potential positive impacts such an/that implementation could bring about and not from their beliefs about the readiness of the people and resources involved in the process. This suggests the university really needed to work harder to better support and facilitate staff and students during the implementation process.

The data obtained from the interview sessions accorded with the questionnaire data. Of eight teachers interviewed, four strongly supported the need to apply the CEFR to non-English major students; three acknowledged the need but held concerns and reservations and one did not think it necessary to implement the CEFR. Supportive ideas yielded from the interview sessions were as follow: Firstly, the division by CEFR of language proficiency into six attainment levels made it more appropriate for different groups of language learners. For non-English major students, applying the CEFR-aligned outcomes of A1 and B1 seemed to be more practical and appropriate compared with previous standards, which were closely aligned with TOEIC and TOEFL tests. One teacher further explained that previous standards were more academic and thus more challenging for non-English major students whose language needs should be more focused on daily and communicative needs. This is understandable because the A1 and B1 CEFR descriptors are mainly focused on "familiar matters regularly encountered in work, school, leisure, etc." (Cambridge 2011: 24), making them more appropriate for non-English major students.

Reasons for teachers' support also came from the expectation that CEFR implementation could create big changes to their teaching and learning contexts, either for the short or long term. In particular, one teacher mentioned the change in students' awareness which led to the changes in "learning methodology". Another added that "the policy is a motivation for students' language improvement". One teacher reflected, "it [the CEFR] affects students' perceptions, which (hopefully) will result in changing students' language competency". All interviewed teachers acknowledged the change in students' attitudes and motivation, which they confirmed to be present and easily recognized in their classes. Nevertheless, they were reluctant to discuss the actual changes in students' language competency and proficiency and admitted that such expectations were "too ambitious" to achieve, even six years after CEFR implementation began in Vietnam.

The second change pinpointed by all eight teachers was the modification and adaptation to teaching practices teachers had made, whether or not done voluntarily. They mentioned what they had done in their classes as evidence of their efforts to make changes accommodating the new policy and implementation. In short, the interviewed teachers observed three additional and direct impacts of the CEFR implementation: *changing students' attitude and motivation in English learning*, *improving teachers' classroom practices* and to some extent *improving the university's qualifications and reputation* and gave these as essential reasons for applying the CEFR framework to their non-English major students.

For teachers who did not perceive the CEFR implementation as necessary, doubt about its efficiency was the main reason given. They pointed to some previous standard-based outcomes and curricula as examples of unsuccessful policies and doubted that the CEFR implementation policy would fare any better. One teacher mentioned suitable planning and reasonable timelines as two basic principles for the students to achieve B1 level. In her view, these two key things were missing from the current environment of Hue University. Reluctance to change and adaptation to changes were additional

reasons for teachers' disagreeing with the requirement to implement CEFR. These teachers expressed their weariness at the previously abrupt and uninformed changes in language policy, specifically to the B1 standard-based learning outcomes, being unexpectedly imposed on teachers and students with limited notice and preparation time. They also expressed fear that just when they became accustomed to a new policy, the policy changed, making them, as one teacher stated: "passive and under a lot of unnecessary pressure". In short, although these concerns and disagreements were not prominent, they helped explaining why GE teachers did not consider the necessity to implement the CEFR as being high; ranking it the lowest average mean score of the four clusters.

3.1.3 GE teachers' perceptions of the CEFR readiness for application

In general, teachers partly agreed that the CEFR and its descriptors applied well to non-English major students, showing an average mean value of 3.71 for this cluster of questions. The mean value of individual items, however, varied greatly from a low of 3.33 to a high of 4.17. Specifically, GE teachers strongly believed that the descriptions of the CEFR levels of proficiency are representative (M=4.06) on the one hand, and that the CEFR needs to be more specific (M=4.17) on the other. Doubts that the descriptors are context-specific or English specific still remained but were not as strong (M=3.33 and 3.39 respectively).

The high SD values of nearly 1.0 to a majority of items showed that teachers' choices were dispersed, indicating inconsistency between individual teacher's perceptions of CEFR specificity. Given that the CEFR descriptors are neither language- nor context-specific, with the descriptions used for each level of proficiency being illustrative rather than representative (CoE 2001). This result should be given serious consideration. The teachers need better understanding of the levels of comprehensiveness of the CEFR descriptors as to use them more effectively.

The data from the interview sessions further explained teachers' perceptions and provided reasons for the quantitative results above. From the interviews, the contradiction between teachers' thinking could be identified and explained. On the one hand, teachers seemed to correctly understand that the CEFR is not a precise document that can be readily applied in every context without modification or adaptation. On the other hand, they were initially hesitant to talk about their uneasiness with the CEFR, which aspects of the CEFR are not suitable and which need improvement to make them more useable or relevant. This might be partly because they were not well trained in understanding this at the outset so did not feel confident enough to say what they think, and partly because of their commonly expressed belief that, as a global framework, the CEFR must be good and complete. Only after encouragement did the participants reveal their concerns more openly and completely. These concerns are described below.

Firstly, four of the eight interviewed teachers strongly agreed that the CEFR descriptors were representative and comprehensive in the levels of proficiency they seek to describe. The main reasons given were that language use at each level was not only divided into skills and sub-skills but also into domains, situations, areas, topics and strategies with all being clearly described for each proficiency level. On the CEFR implementation for non-English major students, however, the teachers provided detailed examples of the inappropriateness of the CEFR descriptors. Some of the descriptors were described as being *alienated* from Vietnamese students' age, ability, interest and concerns. They were also criticized for being *not specific*. The way terms like "basic", "short", "simple", "satisfactory" were used to describe levels of language proficiency failed to help teachers and students visualize clearly the scope and boundary of different levels. This finding accorded with warnings of CEFR limitations pointed out by Little (2006), showing limits to teachers' sound understanding of the CEFR and its descriptors. In addition, the finding was similar to that of Pham (2017). GE teachers also provided evidence of the mismatch between the CEFR and the current context of implementation, due to students' cultural differences, the reality of language need and students' level of proficiency.

3.1.4 GE teachers' dissatisfaction of the work involved in the CEFR implementation process

As seen in Table 3, the low average mean value of 3.19 for the whole cluster, close to point 3 of the five-point Likert scale, showed that teachers were far from satisfied with what had been done to implement the CEFR for non-English major students at Hue University. While some actions were acknowledged, others received strong criticism from the GE teachers, reflected in the wide range, from 3.86 to 1.56, of mean values between items. In particular, GE teachers agreed with the proposition that necessary resources and capacity building for the CEFR implementation had been provided. The mean values for the two items were 3.86 and 3.81 respectively. While GE teachers reported that they were trained, the training and workshops provided the teachers with knowledge of the CEFR's value (M= 3.67) rather than preparing them to apply the procedures (M= 3.39). Results from the questionnaire showed teachers had a neutral attitude towards the feasibility of the timeline (M= 3.06). In contrast, the last three items regarding the available support from experts, the piloting phase of the program and the involvement of teachers and students in CEFR-aligned curriculum design received negative comments from teachers, with all mean values below level 3 (2.69, 2.56 and 1.56 respectively).

Findings from the interview sessions provided better understanding of the data derived from the questionnaire. Although varying in number, all GE teachers interviewed reported their participation in workshops and training, organized by either MOET or their home university, related to the CEFR, its values and limitations and its descriptors. They observed and rated the facilities and resources made available for the CEFR implementation process. Better-equipped classrooms with computers, projectors, CD-players, together with supportive online software and programs were among resources listed by respondent teachers as efforts made by the university to help teachers and students achieve B1 level as the new standard-based learning outcome. They also listed their retraining and improving language proficiency workshops and the English proficiency tests that they participated in from 2011 to 2013 as evidence of the capacity building the university had provided in preparation for implementation. However, all teachers asserted that the CEFR-aligned curriculum was not piloted and they had no significant involvement in its design and development. It can be seen that, while the teachers had relatively sound understanding and perceptions of the CEFR, they were not well prepared for the process of actually implementing it in their own university context.

The interview data revealed that GE teachers were dissatisfied with the implementation process. Their discontent is associated with three main issues, namely time constraints, incompatible teaching materials and the tremendous gaps between students' entry levels of English proficiency and meeting the B1 learning outcome.

3.1.5 Time constraints

In interviews, GE teachers reported their dissatisfaction with the limited number of teacher-led hours assigned to each course. This was the biggest disquiet for GE teachers and led to the two other discontents. The phrase "time constraints" was repeated many times during six teacher interviews. In fact, for non-English major students at Hue University the curriculum specifies 30 teacher-led hours for A1 and A2 courses and 45 hours for B1 courses, which was stated to be "too limited do to anything".

One teacher complained: "We need adequate time to change students' language competence. Yet time allowance [for my non-English major students] to move from A1 to B1 is too limited". This viewpoint was shared by another teacher with her reflection that "the total 30 or 45 periods are not enough to improve students' language proficiency". The phrase "the pressure of time limits" was also raised in other teachers' interviews.

Limited, teacher-led, classroom interactions per week was another cause of the dissatisfaction expressed around time constraints. Due to the limit of 30 or 45 hours, non-English major students at Hue University attended only one class of two or three teacher-led hours each week. "The long interval

between one English classes and the next is enough for my students to forget everything (about English)", one teacher said.

A senior teacher with more than 25 years of teaching experience reported that time allowances for English language curricula for non-English major students had once been much longer, when the school-year programme was applied. The shift from a school-year to a credit-based programme considerably reduced the number of teacher-led, or classroom contact hours while increasing the time allotted to student self-study (or study outside the classroom without a teacher). For language learning, especially for non-English major students, this model has created huge challenges: "simply because not many non-English major students want and have the ability to self-study".

In short, with the current CEFR-aligned outcomes, insufficient time allowance was the biggest pressure GE teachers currently had to deal with. This finding is similar to what Faez, et al. (2011) found in their study where teachers indicated "time crunch" and insufficient time to implement CEFR activities and cover the demanding curriculum simultaneously.

3.1.6 Incompatible teaching materials

The dissatisfaction with the CEFR implementation process, reported by many teachers, was the mismatch between the assigned textbook and the CEFR-aligned outcomes. Many teachers noted that, together with the implementation of the CEFR-aligned outcomes, a new textbook series, *English Elements*, plus a later text entitled *Life,* were selected for course use by non-English major students at Hue University. *English Elements* was severely criticized as being incompatible with the CEFR-aligned outcomes. Some complaints and criticisms are cited below.

Many teachers maintained that *English Elements,* a textbook series by German publisher Hueber, was intended for and targeted on learners who were very unlike students at Hue University. In addition, teachers stated that the series was totally unsuited to the needs of a 105-period English curriculum. Selecting this series for non-English major students at Hue University caused challenges for both teachers and students. As one teacher explained:

> It's impossible to teach four books from the series [*English Elements*] in 105 periods, spread over a total of three semesters. Yet we had to. Comparing the CEFR descriptors for A1-B1 levels, we found that the books contained many irrelevant topics and themes, irrelevant exercises, irrelevant vocabulary and grammar….Some [vocabulary, grammar, topics, etc.] reappears or are repeated in more than one book, while many others, included in the descriptors, cannot be found anywhere [in the textbooks].

Regarding the textbook series *Life*, four (4) teachers reported that this series was better aligned with the A1-B1 CEFR learning outcomes as it focused more equally on the four basic language skills. However, its design indicated that its use required far longer than the 105 periods allocated in the current curriculum. Although challenges arose less from the book itself, GE teachers described problems in selecting content that would help students achieving the required learning outcomes within the allotted time. A senior teacher explained the problems with *Life* as follows:

> Take the A1 course as an example. Each unit in Life has six parts, from A to F, and a review, usually 12 pages long. And we have to teach 6 units, plus administer a mid-term test and an end-of-course speaking test. To do all this we have four periods per unit and three book pages per period. It is too challenging really.

In short, for the CEFR implementation process to be successful and to create changes, GE teachers needed to put in a lot of effort to develop and modify the text materials to align them with CEFR learning outcomes.

This demonstrates that, when the MOET set the CEFR B1 level of proficiency as the required learning outcome, teachers expected that the materials selected should support the achievement of this outcome. It also suggests their belief in the existence of suitable, ready-to-use materials. In contrast, however, teacher feedback on the text materials themselves showed a greater concern with how to deliver the materials within the limited timeframe rather than on how to make effective use of the prescribed materials. They showed less concern to evaluate the materials, adapt and prioritize sections, or select the tasks and topics most useful in supporting student acquisition of the required B1 level of proficiency than for the time limits imposed.

3.1.6 Mismatch between students' admission level of proficiency and learning outcomes

The third dissatisfaction originates from low levels of students' language proficiency at the course entry point. Two teachers thought that students' current proficiency was too low to allow them to achieve the B1 outcome (level three of the six levels) required of non-English major students after three semesters of university study. They cited the low percentage of non-English major students achieving the B1 certificate as evidence of this viewpoint. Six teachers mentioned the vast gap between students' actual English language competency and the level they were required to reach. It was also observed that the situation varied between students undertaking different majors and attending different colleges. One teacher commented:

> It depends on the students. In general, GE students majoring in medicine, pharmacy, or economics have better English language competency compared with students completing majors in other subjects. The B1-aligned outcome may be ok for them, if those students keep on working on their English. But the others, who form the majority, are not good enough.

This idea was widely held, with another teacher stating:

> We did have a placement test before admission so that we could classify students into different ability groups based on their level of English proficiency at entry. I would say that there are many students whose English was at A0 or lower. They simply knew nothing about English despite spending up to ten years learning English at primary, secondary and high schools. How can their English reach B1 level after 105 periods at Hue University?

In conclusion, although the problems may not come directly from the CEFR and the policy to implement it, the reality is that the low levels of students' English ability at the point of course entry have created huge challenges for both teachers and non-English major students at Hue University. From the viewpoint of those having to implement the policy, the mismatch between students' entry levels of English language proficiency and the standard they are required to achieve means that the outcome of students attaining a CEFR level B1 is totally unrealistic.

4 Conclusions and implications

The present study reveals some interesting findings regarding GE teachers' perceptions of the value and the necessity of applying a CEFR-aligned curriculum with standard-based learning outcomes in a specific context. It also displays their attitudes towards its implementation at the grass roots or classroom level. As "change in education is easy to propose, hard to implement and extraordinarily difficult to sustain" (Hargreaves and Fink 2006: 6), some implications and suggestions have been drawn.

Teachers' sound understanding of the value of CEFR coupled with their awareness of the requirement to implement the program within their university can be interpreted as willingness on their part to accept change and innovation in their classrooms, allowing a process whereby "perceptions influence practices" (Borg 2009). However, as a counterbalance, the study also shows that when it comes to the implementation process, GE teachers were not well prepared. Their needs were around lack of resources

and an understanding of the realities they were faced with. Their doubts about achieving positive results from such a program arose from a number of practical factors which together detracted from achieving the required CEFR outcomes. Given that change and innovation take place only when teachers perceive them as feasible (Van den Branden 2009), the GE teachers needed to be given a better understanding of how the changes would occur, what would be involved, and what practical problems to expect during the process. They need access to a forum where they can raise voices and make suggestions around the implementation process. The findings of this study also show that further studies should be conducted especially on teachers' actual practice as response to the implementation of such a global framework as the CEFR.

5 References

Alderson, J. Charles. (ed.). 2002. *Common European Framework of Reference for Languages: Learning, teaching, assessment: Case studies*. Strasbourg, France: Council of Europe.

Bonnet, Gérard. 2007. The CEFR and education policies in Europe. *The Modern Language Journal* 91(4). 669-672.

Borg, Simon. 2009. Introducing language teacher cognition. http://www.education.leeds.ac.uk/research/files/145.pdf (accessed July 26th 2017).

Byram, Michael & Lynne Parmenter (eds.). 2012. *The Common European Framework of Reference: The globalisation of language education policy* (Vol. 23). Bristol: Multilingual matters.

Byrnes, Heidi. 2007. Developing national language education policies: Reflections on the CEFR. *The Modern Language Journal* 91(4). 679-685.

Cohen, David K. & Ball, Deborah Loewenberg. 1990. Relations between policy and practice: A commentary. *Educational Evaluation and Policy Analysis* 12(3). 331-338.

Council of Europe. 2001. *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.

Cambridge, ESOL. 2011. *Using the CEFR: Principles of good practice*. Cambridge: Cambridge ESOL.

Creswell, John W. 1998. *Qualitative inquiry and research design: Choosing among five traditions.* Thousand Oaks, CA: Sage.

Creswell, John W. 2013. *Research design: Qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, CA: Sage.

Creswell, John W. & Vicki L. Plano Clark. 2007. *Designing and conducting mixed methods research.* Thousand Oaks, CA: Sage.

Faez, Farahnaz, Suzanne Majhanovich, Shelley Taylor, Maureen Smith & Kelly Crowley. 2011. The power of "Can Do" statements: teachers' perceptions of CEFR-informed instruction in French as a second language classrooms in Ontario. *Canadian Journal of Applied Linguistics* 14(2). 1-19.

Hargreaves, Andy & Dean Fink. 2006. *Sustainable leadership*. San Francisco, CA: Jossey-Bass.

Hulstijn, Jan H. 2007. The shaky ground beneath the CEFR: Quantitative and qualitative dimensions of language Proficiency. *The Modern Language Journal* 91(4). 663-667.

Kaplan, Robert B. & Richard B. Baldauf. 1997. *Language planning from practice to theory* (Vol. 108). Clevedon: Multilingual Matters.

Kennedy, Chris. 2013. Models of change and innovation. In Ken Hyland & Lillian L. C. Wong (eds.), *Innovation and change in English language education*, 13-27. Abington: Routledge.

Little, David. 2006. The Common European Framework of Reference for Languages: Content, purpose, origin, reception and impact. *Language Teaching* 39(3).167–190.

Little, David. 2007. The Common European Framework of Reference for Languages: Perspectives on the making of supranational language education policy. *The Modern Language Journal* 91(4). 645-655.

McLaughlin, Milbrey Wallin. 1987. Learning from experience: Lessons from policy implementation. *Educational Evaluation and Policy Analysis* 9(2).171-178.

Ministry of Education and Training (MOET). 2008. Decision No. 1400/QD-TTg Approval of the Project "Teaching and learning foreign languages in the national educational system for the 2008-2020 period". Hanoi, September, 2008.

MOET. 2014. Circular No. 01/TT-BGDĐT Issuing Six-level framework for foreign language proficiency in Vietnam. Hanoi, January, 2014.

MOET. 2014. 2020 Project Report. *TESOL symposium.* Danang, Danang University.

Nguyen, Van Huy & M. Obaidul Hamid. 2015. Educational policy borrowing in a globalized world: A case study of Common European Framework of Reference for Languages in a Vietnamese university. *English Teaching: Practice & Critique* 14(1). 60-74.

Pham, Thi Hong Nhung. 2012. Applying the CEFR to the teaching and learning English in Vietnam: Advantages and challenges. *Journal of Foreign Language Studies* 30. 90-102.

Pham, Thi Hong Nhung. 2017. Chapter 6. Applying the CEFR to renew a general English curriculum: Successes, remaining issues and lessons from Vietnam. In Fergus O'Dwyer et al. (eds.), *Critical, constructive assessment of CEFR-informed language teaching in Japan and beyond*, 97-117. Cambridge: Cambridge University Press.

Pham, Thi Hong Nhung. 2018. General English Proficiency or English for Teaching? The Preferences of In-service Teachers. *RELC Journal* 49(3). 339–352.

Poon, Anita Y. (2000). *Medium of instruction in Hong Kong: Policy and practice*. Lanham: University Press of America.

Steiner-Khamsi, Gita. (ed.). 2004. *The global politics of educational borrowing and lending*. New York: Teachers College Press.

Tono, Yukio & Masashi Negishi. 2012. The CEFR-J: Adapting the CEFR for English language teaching in Japan. *Framework & Language Portfolio SIG Newsletter* 8. 5-12.

Van den Branden, K. 2009. Diffusion and implementation of innovations. In Michael H. Long & Catherine J. Doughty (eds.), *The handbook of language teaching*, 659-672. Oxford: Blackwell Publishers.

Vietnamese government. 2017. Decision No. 2080/QĐ-TTg dated 22/12/2017 Approval of the expansion of the National project "Teaching and learning foreign languages in the national educational system for the 2008-2025 period". Hanoi, December, 2017.

Wang, Hong. 2008. Language policy implementation: A look at teachers' perceptions. *Asian EFL Journal* 30(1). 1-38.

Waters, Alan. 2009. Managing innovation in English language education *Language Teaching* 42(4). 421-458.

6 Appendices

Appendix 1. Questionnaire

Respondent's code: ____

Part 1. Personal information

Please tick or write the answers in the squares given.

1. Gender: ☐ male   ☐ female

2. How long have you been teaching non-English major students?

    ☐ 1-5   ☐ 6-10   ☐ 11-20   ☐ more than 20 years

3. What is your highest qualification?

    ☐ Bachelor   ☐ Master   ☐ Doctor (PhD)

4. Have you got another Bachelor Degree beside English one?   □ Yes   □ No

5. Whose workshops on CEFR have you attended?

    □   By MOET

    □   By home university

    □   Others: _____

Part 2. The implementation of the CEFR at your university

Please circle the number reflecting the level of your agreement.

5: strongly agree, 4: agree; 3: neutral; 2: disagree; 1: strongly disagree

No.	Statements	5	4	3	2	1
1.	Necessary resources for the implementation were provided.	5	4	3	2	1
2.	The CEFR-aligned descriptors are representative for the language proficiency of its level.	5	4	3	2	1
3.	The CEFR can make language learning outcomes transparent.	5	4	3	2	1
4.	The implementation of the CEFR was piloted.	5	4	3	2	1
5.	The CEFR allows mutual recognition across institutions.	5	4	3	2	1
6.	Capacity building for the implementation (e.g. training workshops on the CEFR) was provided.	5	4	3	2	1
7.	Staff involved was informed about the values and limitations of the CEFR.	5	4	3	2	1
8.	The CEFR is meant to encourage self-directed language learning.	5	4	3	2	1
9.	The CEFR is applicable because it is English-specific.	5	4	3	2	1
10.	Teachers were involved in the CEFR-aligned curriculum design.	5	4	3	2	1
11.	The CEFR is applicable because it is context-specific.	5	4	3	2	1
12.	The CEFR can be used as a basis for the renewal of classroom assessment.	5	4	3	2	1
13.	The CEFR can be used as a basis for the renewal of the language teaching curriculum.	5	4	3	2	1
14.	Staff involved was trained for the implementation procedure.	5	4	3	2	1
15.	The CEFR is ready for any curriculum renewal.	5	4	3	2	1
16.	Expertise and professional support during the implementation process were provided.	5	4	3	2	1
17.	The CEFR-aligned descriptors need to be further specified to be applicable to the context in which it is used.	5	4	3	2	1
18.	The CEFR can be used for positive change in English language education.	5	4	3	2	1
19.	The objectives were realistic within the required timeline.	5	4	3	2	1

20.	The present implementation of the CEFR in Vietnam is necessary as:					
	CEFR is a global comprehensive framework.	5	4	3	2	1
	Teachers involved in the process are ready.	5	4	3	2	1
	Students involved are ready for such an application.	5	4	3	2	1
	CEFR has been well applied in many other countries for innovations in language teaching.	5	4	3	2	1
	My university has all the resources required for such an application.	5	4	3	2	1
	CEFR can help improve the teaching quality of the university.	5	4	3	2	1
	CEFR can help my university promote its reputation.	5	4	3	2	1
	CEFR implementation will improve the language proficiency of the students of the university.	5	4	3	2	1

Thank you for your cooperation.

Appendix 2. Main Interview Questions

(translated from the original Vietnamese version)

1. What do you know about the CEFR?
2. In your opinion why the CEFR is adopted at your university?
3. What do you know about the decision-making processes of applying the CEFR for non-English major students at your university?
4. Is the implementation of the CEFR for non-English major students at your university necessary? In what ways?
5. Do you think that the CEFR is ready for implementation for non-English major students at your home university? In what ways?
6. What do you think about the implementation of the CEFR for non-English major students at your home university? And why?
7. What challenges and difficulties have you encountered so far due to the CEFR implementation? What are the reasons for these problems?
8. What are your suggestions for effective implementation of the CEFR at your university and in contexts alike?

7 Biographies

Le Thi Thanh Hai is a lecturer of the Department of English for Specific Purposes of the University of Foreign Languages, Hue University. She has been teaching English for 19 years and involved in different projects on developing training curriculum for in-service English language teachers. She is a currently a PhD student of Hue University of Foreign Languages.

Pham Thi Hong Nhung is a senior lecturer of the Department of English, University of Foreign Languages, Hue University. She obtained a Master degree and a PhD degree in Applied linguistics from the University of Queensland, Australia. Pham has 22 years of experience teaching and training in-service English language teachers. She has lead various national research projects. Her research interests include language policy, teacher education, and intercultural communication.

Jumping through hoops and keeping the human-in-the-loop
—Interview with Dr Nick Saville

At the JALT International Conference in Tsukuba, Japan, in November 2017, two of the *CEFR Journal*'s editors delighted in being offered the opportunity to interview Dr Nick Saville (Director of Research and Thought Leadership at Cambridge Assessment English). At the conference, Dr Saville presented a keynote speech entitled *Data & Devices: the 4th Industrial Revolution & Learning*, as well as a workshop focusing on *LOA: Understanding & Using Assessment to Support Learning*. LOA here stands for *Learning Oriented Assessment*.

In the interview, we were hoping to elicit some insights and answers about dealing with technology in language learning, teaching, and assessment, as well as on issues related to the CEFR in general, and the Cambridge English Profile[1] series in particular. Towards the end of the interview, we asked some self-referential questions. To make our intentions very clear, we were hoping they might further aid our readers in understanding what we are aiming to achieve by launching this journal. To whom is the journal addressed? And, most importantly, why are we seeing a need for such a journal to fill a space previously sparsely filled at best? We would be delighted were the kind reader to overlook this insolence and not mistake it for improper indulgence or undue navel gazing. Thank you.

Keywords: Japan, CEFR-J, assessment, artificial intelligence, Cambridge Maxims, productive skills

Morten Hunke: Thank you for agreeing to talk to us and answer a few questions! Let's see if we can get through most of them, but of course we have a bit of a priority list. I'll just start with an open question on the CEFR: what was your first contact with the CEFR? What do you think are the strengths and the weaknesses of the framework?

Dr Nick Saville: I trace the CEFR back to the early '70s (1970-1972) when the modern languages project of the Council of Europe (CoE) started. It was an evolution of *learning objectives* and *levels* which started at that time.

My first interaction with the CoE level system was in 1987, when I first came to Japan. At the time, I used the Waystage and Threshold[2] (van Ek and Trim 1991a, 1991b) levels to inform the specifications for two tests—the Pre-PET and the PET (Preliminary English Test)—which are now A2 Key[3] and B1 Preliminary respectively[4] in the Cambridge English Qualifications, named for their CEFR level.

Interestingly, both those tests were designed and evolved from the Japanese context, when I was working here with local partners to introduce a more communicative approach. It was based on the understanding at the time of communication and levels, which were the forerunner of the CEFR that finally came out in 2001.

In addition to that, in 1990, Cambridge University Press (CUP) published the revision of the Waystage and Threshold levels and it was around that time that I began my association with John Trim. Also, independently of the CoE, a group of test providers set up ALTE[5], the Association of Language Testers

---

1. English Profile: https://www.englishprofile.org/?
2. T-series Books: https://www.englishprofile.org/component/content/article?id=119&It=.
3. A2 Key: https://www.cambridgeenglish.org/exams-and-tests/key/
4. B1 Preliminary: https://www.cambridgeenglish.org/exams-and-tests/preliminary/
5. Association of Language Testers in Europe (ALTE): https://www.alte.org/.

in Europe, and published a 5-level system in 1991 that incorporated the Waystage at Level 1. Level 2 was the Threshold Level, and in the course of the 1990s, ALTE added in a Breakthrough Level (now A1) through a project carried out by the so-called FINGS Group—the Finnish, Irish, Norwegian, German and Swedish members—that had an interest in developing this lower level.

So ALTE was developing a 5- or 6-level system (depending on how you think of it) at the same time as the Rüschlikon Conference in 1991 (Little, Gollier and Hughes 2011). This conference was the impetus for developing the CEFR as we now know it—as well as the European Language Portfolio Project.

So through Cambridge and ALTE, I have been working on the CEFR concept since that time. My colleague Michael Milanovic, who was then the manager of ALTE, represented Cambridge and ALTE on the Sounding Board Group. This was a group of invited experts that helped the authors and the CoE put together the CEFR. I remember going to the launch of the pre-publication version in 1996 in Strasbourg as an expert in a wider consultation group of experts.

In the period between 1996 and the publication date (end of 2000), we interacted with the CoE in various ways to help them collect data about the Pilot Versions. I was on a group that was consulted about the editing of the final document, and ALTE provided one of the appendices. ALTE and DIALANG had both developed Can Do statements in parallel to the CEFR and these were included as additional examples (CoE 2001).

In the earlier publication (dated 1995/96), the Can Do statements that Brian North had validated were seen as exemplars and appeared in the appendix rather than in the body of the text. By 2001, only the ALTE (2002) and DIALANG ones remained as appendices, and the others were incorporated into the body of the CoE text as we currently know it (2001, 2002).

Morten Hunke: According to Little (2011), the CEFR was initially intended to facilitate closer interdependency between curriculum, pedagogy and assessment. He also goes on to suggest working towards an assessment culture in which external tests and exams exist in a continuum with teacher assessment, peer assessment and learner self-assessment. Do you think the four maxims[6] that Cambridge English (2016) stipulate, provide all the stakeholders with opportunities to refer back to the action and reflection principles that underpin the CEFR?

Dr Nick Saville: Well, yes, I agree with David. In fact, his vision of the relationship with learning and teaching assessment has actually evolved in Cambridge English around the notion of Learning Oriented Assessment or LOA, which as you know, I call a *systemic approach*. It is one that brings together all stakeholders to facilitate learning and to allow learners and teachers to demonstrate their skills—it aims to create an effective *ecosystem of learning.*

The Cambridge English maxims for achieving *positive impact by design*—which we designed back in in 1995 or 1996—were aimed at achieving positive impact of our exams in local contexts. So, it is really an approach that enables assessment providers to ensure that stakeholders are consulted and informed, and that we demonstrate through collecting evidence that the intended impacts, including washback and so on, are achieved when implementing a testing project.

Morten Hunke: I see, but what does this mean in practice?

Dr Nick Saville: The idea is that you need a rational and well-planned approach to your assessment design that allows you *ab initio* to come up with an impact-by-design concept. In other words, you have a clear idea at the start of the impact that you hope to achieve.

In implementing your testing project, there's no point coming up with a great test if the majority of the stakeholders fail to understand what the constructs are, or what the intended impacts should be when we put it into practice in context.

Communication with the stakeholders is therefore fantastically important. I think there's typically not enough focus on helping people to understand assessment in their own contexts—what is now known

---

6.   Maxim 1: Plan, Maxim 2: Support, Maxim 3: Communicate, Maxim 4: Monitor and Evaluate

as assessment literacy.

This requires a lot of support from the assessment providers themselves—an infrastructure to provide relevant materials, to train stakeholders, to 'hold their hand' if you like, to check back that they 'get it'. This 'rational model' is an iterative one, where you are collecting information and adjusting things as you go along.

Of course, monitoring doesn't just stop at a certain point; it's a cyclical process. So, after five years, for example, you need to know whether the effects and consequences are still the ones that you observed at the beginning. Or if you discern that things don't go as planned, or there are unexpected or unintended consequences—which is typical in most social and educational behaviours—you must be able to adjust your system to respond to this evidence. As a responsible body involved in implementing educational reforms, you cannot carry on regardless. That is a recipe for negative impact.

Morten Hunke: So, that includes you trying to have as much communication as possible with the test centres, people who are facilitators in the countries who are then talking to the examiners-to-be and the current examiners?

Dr Nick Saville: Yes. I mean it's a systemic, networking approach. You can't do exams *to* people; exams and other forms of assessment are always embedded within an educational context. This means that all forms of assessment need to be both externally valid and locally implemented, so that what is determined to be the outcomes actually get implemented in practice with the stakeholders in the contexts where they are used.

Of course, education reform is always a slow process. Often, it is not longitudinally planned with enough 'runway' to achieve what is needed. You need to have active participation and to take your stakeholders with you. Working side-by-side, assessment providers can help them bring about intended improvements to the educational outcomes.

Morten Hunke: On that note, you said earlier that KET (Key English Test) was 'begotten' in the Japanese context—that it developed from ideas you and other people had while working here in the 1980s.

And someone who runs a test centre here in Japan told me, the one thing that prevents people—especially high school students—from taking the test is that most test centres offer the tests on Saturdays only. Of course, this is a very concrete thing, but it's also a communication thing.

Dr Nick Saville: You're right, it's an issue of making international exams more accessible in local contexts. International assessment providers like Cambridge English tend to have international dates, and this can sometimes be a problem locally.

When we designed Pre-PET for Japan, its early trials and implementations were on Sundays to fit local preferences. At that time, we could fit in with the Japanese school system. Later the Pre-PET model was adapted as an international test and it became Key (or KET) in 1994. The Sunday dates were suspended as they are not popular in most countries.

Going forward, I think one of the things that is going to change is that technology and computer-delivered tests will enable more frequent dates and greater flexibility to meet local needs. This is what we're increasingly attempting to do—to customise and personalise our service to respond to local requirements. This means becoming increasingly 'learner-centric'.

I think over the next few years you will see more Cambridge English centres opening up in Japan, and a growing interest amongst Japanese learners in taking international tests beyond the ones that are currently available.

Thirty years ago, EIKEN[7] was particularly strong and their tests were embedded in the school system—I imagine this is still the case to some extent? With the test dates published in staff rooms and with the teachers being EIKEN-trained examiners, it meant that the EIKEN approach actually became part of the education system. Similarly, in the world of work, TOEIC[8] has occupied that space for more than 30 years.

---

7.   Eiken Foundation of Japan: http://www.eiken.or.jp/eiken/en/

8.   TOEIC: https://www.ets.org/toeic

It seems that there's now an appetite for alternative approaches and more choice—and there is an impetus from the Japanese government to promote four-skills tests as ways into university. This will mean that the Cambridge English learning-oriented approach may become both more relevant in the local context, and more widely recognised. This will give it the currency needed to make it worth the effort to prepare for our exams.

Morten Hunke: What about the communicative mission the CEFR has? That very much permeates the Cambridge English tests. How does that relate to the simultaneous use of technology?

How do you think using technology for rating—writing and/or speaking—can actually be something that sends us on a path towards the future? What role do human raters still have to play in a scenario where artificial intelligence (AI) is rating writing and speaking, and potentially other parts of tests as well?

Dr Nick Saville: That's a very good question. But I think there's a way to go before humans are not part of the equation anymore.

At the moment, in Cambridge English we don't see our auto-rating system as a replacement for humans. Currently IELTS[9]—which, as you probably know, is the biggest international examination for academic proficiency testing, with more than three million candidates—still has an obligatory face-to-face oral test. And actually, it's one of the things the learners like best. Some test takers have a negative attitude towards current speaking tests based on 'talking to a computer'. This may be because current computer-based speaking tests do not provide interactive communication. I think it can feel very similar to 'talking to the wall', not dissimilar to the experience of using the old kind of language lab we had in the 1970s. It's not very motivating and it's not very much like real conversation.

However, the combination of humans and machines—what I call the 'virtuous combination' of what the human can do backed up by what the machine can do—is how I see AI going in Cambridge English. In other words, we aim to put humans and computers together to get improved benefits for both learning and assessment.

We already have a tool for assessing writing which works very well for low-stakes testing in learning contexts—it is known as Write & Improve[10]. The AI tool which underpins it is an auto-rater which can rate learners' writing as accurately as human raters for this purpose, and can provide learning-oriented feedback as well. This makes it very useful for use in certain contexts—for the lower-impact decisions that you might want to make, for institutional purposes, placement testing, benchmark testing and so on.

If you ramp up the stakes, your assessment system needs to be increasingly dependable. If you lower them, the outcomes can be mitigated by other decisions or contextual features. I think at the lower end of this continuum, you can have machine-delivered assessments, and in Cambridge English we are already there for assessing writing, and we're almost there for speaking. We are not yet there for high-stakes purposes—and we might never be if we decide that it is desirable to 'keep the human in the loop'.

Morten Hunke: I'm aware that Cambridge English is quite adamant in trying to have good communication with the people in the countries, and not only to communicate, but rather to factor in whatever special scenarios and situations exist in the country. In Japan, there is a situation where multiple-choice questions (MCQs) are king, so is it possible to marry aiming at testing productive competence and having largely MCQs? Is that something that is at all possible or would you say this is going against LOA?

Dr Nick Saville: If you're trying to test the unobservable—something going on in your brain such as the ability to understand texts through reading—you currently need to do this in an indirect way. You can't yet observe the thinking process by putting electrodes on peoples' heads or through other kinds of clever technology. Perhaps for the future?

Currently you have to elicit a response to infer how people are understanding things. In carefully designed tests of reading, for example, MCQs—particularly in a task-based context—work quite well.

9.   IELTS: https://www.ielts.org/

10.   Cambridge English Write & Improve: https://writeandimprove.com/

But I don't think there should be a role for discrete-point multiple-choice grammar items which are atomistic—items which are decontextualized and which can be crammed for. They don't fulfil my idea of task-based assessment and they don't fulfil the idea of something which generates the cognition akin to using the language in the target use situation. They are effectively test behaviours that can be taught as a surrogate for learning the language. So yes, I think there is a place for task-based MCQs in a communicative environment and with learning at the heart for the receptive skills.

Morten Hunke: What about productive skills or integrated tasks?

Dr Nick Saville: I think if you're talking about the productive skills, then the only way you are going to be able to test writing and speaking effectively is by getting people to speak and write, or potentially see them as integrated with other skills, i.e. 'read this and speak about it' or 'read this and write about it'.

We talk nowadays about 'six skills': reading/writing, listening/speaking, plus interaction, and mediation. Increasingly, it is seen as construct-relevant to integrate these skills to reflect real-world uses of language. For example if you were trying to recreate an academic environment, consider the following kind of scenario: 'a tutor instructs her students to read three books before their tutorial the following week, and to be prepared to talk about certain key concepts during the seminar before writing an assigned essay for assessment purposes'. How much of a student's participation in the seminar is determined by reading comprehension or fluency in speaking? We don't yet have a construct that easily accounts for this in assessment, but I think it's coming. Such constructs will be easier to operationalise using technology rather than in the traditional paper-and-pencil mode.

Morten Hunke: Coming back a little bit to the electronic online testing format, do you think there are any ethical issues in storing vast amounts of data that could be used—it could be text, it could be audio recordings—is there anything that you're concerned about?

Dr Nick Saville: Personally, I've been concerned about this for 25 years. Data protection laws have been in place at the European level, at national level and at institutional level through all kinds of codes of ethics and legal parameters. You can't keep or store people's electronic data even now without their permission, and you have to make certain things available to people if you store them. In the UK we have very strict data protection laws and even stricter European (GPDR) regulations [have] come into force in 2018.

I think people who work with us can be sure that their data is treated appropriately. It's axiomatic of educational assessment that test takers (or their guardians) must give permission for personal data to be used to make judgements and decisions about individuals. It's part of the contract of doing an international test that you sign up to giving your data to the assessment provider to make a judgment about you.

In the contract that people sign, they can also agree to their data being used for a number of legitimate purposes, such as research and validation. This will be governed by data protection regulations which are legally and institutionally validated, and may include anonymity in the way the data is stored. In Cambridge, we have built a 60 million-word corpus of learner language, taken from learners' writing—with their permission of course—and stored in such a way that meets the requirements of access and control.

As we move forward in the AI world of 'data and devices', the amount of data that will become available will make such corpora appear very small. The ethical and regulatory issues will however become increasingly complex and we will need to be more vigilant about the potential misuse of personal data.

Morten Hunke: Another question related to digitisation and especially to AI: there's a lot of talk about washback effect of tests and, as you have suggested, we know that some pretty negative washback can occur. Do you think using AI can have positive washback effects on learner autonomy—something that is especially key to the CEFR as well—and if so, how?

Dr Nick Saville: I think that the example which I'm talking about at the JALT conference is a very good example of how to implement AI in an ethical way, in order to provide learners with autonomy to learn and to get information and feedback which can help their learning progress. This information can also then can be incorporated into learning programmes.

Cambridge English's Write & Improve is basically what is called a business-to-consumer model; it's mainly aimed at individual learners, but it can also be used by schools and teachers to aggregate information and to be used in a more programmatic way in classroom contexts. It has a feature called Classview for this.

I think you have to strive for transparency and clear explanations about your AI—how it's to be used, what its strengths and limitations are and so on. The aim should be to build trust based on sound principles—both ethical and theoretical—in order to give the public some reassurance that it is not a 'black box' doing things to people without their knowledge, awareness or consent.

There is a growing concern about this. In fact, I heard the CEO of IBM Watson—a big AI programme—talking about this on YouTube at the World Economic Forum (2017). She was saying that AI really needs to be based on these three principles—*trust,* created through *transparency,* and sound *principles*. I think that's what we are trying to do in the field of language education.

Morten Hunke: Moving back a little bit towards the CEFR and the Japanese context in particular—are you aware of research done in Japan into the CEFR and implementations of the CEFR?

Dr Nick Saville: I am indeed. In fact, leading applied linguists here have been looking at the work of the CoE for at least 20 years. I myself have welcomed delegations to Cambridge, including senior professors like Prof Ikuo Koike (Keio and Meikai University) and his associates. I introduced them to Dr John Trim, on more than one occasion in fact.

I've also been to many meetings with Japanese colleagues to discuss the underlying principles of the CEFR. Of course Prof Koike, Prof Tono and Prof Negishi (both Tokyo University of Foreign Studies) directed funded projects to investigate the adaptation of the CEFR to Japan.

The CEFR-J[11], which came out a few years ago, provides a very good example of how the CoE intended the CEFR to be used: i.e. as a document to inform, guide or help people to develop their own implementations without foisting ready-made solutions on people.

The Koike Kaken[12] (research) group spent six or seven years, I think, working through a very rigorous attempt to understand CEFR principles for use in Japan. The group concluded that much of the approach is applicable. But they also realised that the CEFR is not a 'cookie cutter' model. Although the principles provide an impetus for bringing about change, many details needed to be adapted for implementation in Japanese education. For me, it provides perhaps the best example of the CEFR being adapted in an international context.

Maria Gabriela Schmidt: Do you think the CEFR-J is a model case for other countries?

Dr Nick Saville: I think it's an interesting case study, but I don't think a 'model case' is actually needed. That would imply that *this* is the way to do it.

I recently worked with a team in Thailand which decided that they were not going to have a CEFR-Thai, and also with the Chinese Standards of English project. In both cases, the researchers decided that they wanted to use some underlying principles of the CEFR—like the action-oriented, Can Do approach—but that they didn't want to adapt the CEFR in its entirety. That's understandable because the CEFR is actually a vast reference framework and what they wanted was something specifically to guide their national, English language reform programme.

I would say that rather than being a model, CEFR-J is exemplary in the way the team went about their work—providing a rigorous and transparent way of reaching conclusions that have been documented and widely published. Although the CEFR may not be well-known in Japan yet, the publications are there for people to look at it if they need to.

Of course, the latest book in the English Profile Studies (EPS) Series—*Critical, Constructive Assessment of CEFR-informed Language Teaching in Japan and Beyond* (O'Dwyer et al, 2017)— compiles some excellent

---

11. CEFR-J: http://www.cefr-j.org/ – Japanese language only.

12. Please see: https://kaken.nii.ac.jp/

case studies which document how the CEFR-J has begun to have an impact.

EPS volume 6 is a unique collection of papers. It is not uncritical but many positive things are now documented and available for people to reflect on. I think that's the exemplary part.

Morten Hunke: This interview is going to be published in the maiden issue of our new online journal: *CEFR Journal*. This is, at least partially, in reference to Cambridge's own *English Profile Journal.* What do you think are the key aspects required of a journal with a regional focus such as ours is going to have— but dealing with the CEFR—in order to establish itself?

Dr Nick Saville: *English Profile Journal*[13] was really an opportunity when that project was at its height for people who were both working within the project, but also working within a wider network of collaborators, to share information rapidly. Although it was refereed to a high standard, it wasn't designed to set itself up as a rival to some of the applied linguistics or language testing journals, which are highly rated. In that respect, it existed and exists to share information.

Since the English Profile Programme is now in a 'business as usual' phase—it's not pushing hard on any specific topics right now—the input to that journal has rather died down, so it's residing there in the background as a sort of store of information. I think *your* impetus and using the CEFR to have a journal and to raise awareness of the CEFR and the issues in your local or regional context is an excellent idea.

In terms of online journals, I'm working on another project in Cambridge called 'Multilingualism Empowering Individuals and Transforming Society', or MEITS, which is a new interdisciplinary and internationally collaborative project on multilingualism. That project has launched an online journal in different strands.

As a way of getting peer-reviewed articles into the public domain in a timely way in support of the project, i.e. high-quality research or high-quality position papers, it's fantastic. If you wait to get into one of the established journals, then the project funding may be over.

Morten Hunke: Well that's the main idea behind it, not putting just yet another journal out there and having it highly rated. The main aim for the project is to create a forum for people. To allow researchers and practitioners to showcase the things they are doing in a more timely manner, rather than publishing it in a book like EPS volume 6 (O'Dwyer et al 2017). Of course, the volume has interesting case studies and displays a large degree of alignment of all the authors' contributions, but the new journal is really more for the community to actually do what a community is supposed to do: to communicate.

Dr Nick Saville: Yes, but I think it has to be rigorous as well; you've got to set standards, and the community has to accept that if you're not up to the required standard, then you won't get published in the journal. You have to encourage people to 'jump through the hoops', otherwise it might end up being another newsletter or a blog—blogs aren't necessarily low-grade per se, but if you want it as a journal, then you will need to set a higher academic standard. I think that's the main thing.

The fact that things take a long time is a problem with some of the established journals. They only take about 20% of contributions and tend to build up a backlog of papers under review. That's part of setting a very high standard. Getting through the peer review process in a timely way, and then getting revisions done, can mean waiting several years to get into print.

For you, I think the timeliness is really important because it's about sharing ideas. If you can't share a paper in a way that people can react to, it won't function as a sharing tool or as a community-based approach. I am fully in support of you doing it—and in making sure you do it rigorously.

Morten Hunke: That's a very good and interesting point, and in fact the last question I would like to ask is immediately related to this. We intend to be as timely as possible and as absolutely rigorous as we can be, but we are a small group, as you may be aware. What's your impression so far of the work that has been done from within this group, the CEFR & LP SIG[14] (former FLP-SIG)?

---

13. English Profile Journal: https://www.cambridge.org/core/journals/english-profile-journal
14. JALT CEFR & LP SIG: https://cefrjapan.net/

Dr Nick Saville: You've got a great bunch of people here who are doing high-quality work, so I think you have plenty to build on. But if you don't want it to be self-referential, then you might want to find external reviewers, such as an editorial board for the journal, who will be critical and provide a wider perspective.

Maria Gabriela Schmidt/Morten Hunke: Thank you very much!

Dr Nick Saville: Thank you indeed, it has been a very interesting conversation.

References

ALTE. 2002. The ALTE Can Do Project: English version – Articles and Can Do statements produced by the members of ALTE 1992-2002. https://www.cambridgeenglish.org/Images/28906-alte-can-do-document.pdf. (accessed 12 March, 2019.)

Cambridge English. 2016. Principles of Good Practice: Research and innovation in language learning and assessment. https://www.cambridgeenglish.org/Images/22695-principles-of-good-practice.pdf. (accessed 2 March, 2019.)

Council of Europe. 2002. Common European Framework of Reference for Languages: Learning, teaching, assessment – Case Studies.

Council of Europe. 2001. Common European Framework of Reference for Languages: Learning, teaching, assessment. Strasbourg: Council of Europe & Cambridge: Cambridge University Press. http://rm.coe.int/1680459f97. (accessed 12 March, 2019.)

English Profile Journal. Ceased publication. https://www.cambridge.org/core/journals/english-profile-journal. (accessed 2 March, 2019.)

Little, David. 2011. Thinking Allowed. The Common European Framework of Reference for Languages: A research agenda. *Language Teaching* 44(3), 381-393.

Little, David, Francis Gollier & Gareth Hughes. 2011. The European Language Portfolio: the story so far (1991-2011)—Executive Summary. https://www.coe.int/en/web/portfolio/the-european-language-portfolio-the-story-so-far-1991-2011-executive-summary. (accessed 14 March, 2019.)

O'Dwyer, Fergus, Morten Hunke, Alexander Imig, Noriko Nagai, Naoyuki Naganuma & Maria Gabriela Schmidt. 2017. *Critical, Constructive Assessment of CEFR-informed Language Teaching in Japan and Beyond*. Cambridge: Cambridge University Press.

van Ek, Jan A. & John Trim. 1991b. *Waystage Level 1990*. Strasbourg: Council of Europe & Cambridge: Cambridge University Press.

van Ek, Jan A. & John Trim. 1991a. *Threshold Level 1990*. Strasbourg: Council of Europe & Cambridge: Cambridge University Press.

Biography

Dr Nick Saville is Director of the Research and Thought Leadership Division for Cambridge Assessment English. He is Secretary-General of the Association of Language Testers in Europe (ALTE), on the Board of Trustees for The International Research Foundation for English Language Education (TIRF), a Board member of Cambridge University's Institute for Automated Language Teaching and Assessment (ALTA), and a visiting professor for the ICT-Assisted Interpreter Training Project at Xiamen University, China.

He was a founding associate editor of the journal Language Assessment Quarterly and is currently joint editor of the Studies in Language Testing (SiLT, CUP), previously with the late Prof Cyril Weir, and editor of the English Profile Studies series (EPS, CUP). He co-authored a volume on Learning Oriented Assessment (LOA) with Dr Neil Jones (SiLT 45).

CEFR JOURNAL—RESEARCH AND PRACTICE

Submission (Call for Papers)

This journal attempts to fall somewhere in between an inaccessible academic journal (long waiting times, fairly strict guidelines/criteria) and a newsletter (practical in nature but lacking in theoretical support/ foundation), linking research of a practical nature with relevant research related to foreign language education, the CEFR, other language frameworks, and the European Language Portfolio. While the CEFR was introduced by the Council of Europe and intended for use, first and foremost, within Europe, the influence of the CEFR now has to be attested in many places beyond European borders. It has become a global framework, impacting a variety of aspects of language learning, teaching, and assessment across countries and continents beyond the context for which it was originally created. As such, there is a pressing need to create a quality forum for sharing research, experiences, and lessons learned from applying the CEFR in different contexts. This journal provides such a forum where people involved or interested in processes of applying the CEFR can share and learn from one another.

We are continuously seeking contributions related to foreign language education, the CEFR, other language frameworks, and the European Language Portfolio. We are particularly interested in specific contextual adaptations.

Please contact the editors and submit to:

journal@cefrjapan.net

Guidelines

Submission: Annually by March 31st

Contributions: Articles (research), reports (best practice, work in progress, conference presentations), research notes, book reviews, information exchange

Language(s):   English (British, American, international) preferred, but not mandatory.

                    Other languages by request, with an extended abstract in English.

Review type: Peer review, double blind

Peer review guidelines:

We ask all peer reviewers to make every reasonable effort to adhere to the following ethical guidelines for the **CEFR Journal – Research and Practice** submissions that they have agreed to review:

- Reviewers must give unbiased consideration to each manuscript submitted for consideration for publication, and should judge each on its merits, without regard to race, religion, nationality, sex, gender, seniority, or institutional affiliation of the author(s).

- Reviewers should declare any potential conflict of interest prior to agreeing to review a manuscript, including any relationship with the author that may potentially bias their review.

- Reviewers must keep the peer review process confidential; information or correspondence about a manuscript should not be shared with anyone outside the peer review process.

- Reviewers should provide a constructive, comprehensive, evidenced, and appropriately substantial peer review report.

- Reviewers must avoid making statements in their report that might be construed as impugning a person's reputation.

- Reviewers should make all reasonable efforts to submit their report and recommendation in a timely manner, informing the editor if this is not possible.

- Reviewers should call to the journal editor's attention any significant similarity between the manuscript under consideration and any published paper or submitted manuscript of which they are aware.

—Adapted from © Taylor & Francis Online

Author instructions:

- deGruyter Mouton guidelines for Language Learning in Higher Education (CercleS) and style sheet

CEFR Journal – Research and Practice