CEFR Journal

Research and Practice

Japan Association for Language Teaching (JALT)

CEFR & Language Portfolio SIG (CEFR & LP SIG)

Volume 5 (December 2022)

ISSN 2434-849X

Title: CEFR Journal—Research and Practice

Type: Online Journal
URL: https://cefrjapan.net/publications/journal
Contact: journal@cefrjapan.net

Edited by: Japan Association for Language Teaching (JALT)
CEFR & Language Portfolio SIG
Fergus O'Dwyer (editor)
Maria Gabriela Schmidt (coordinator, editor)
Morten Hunke (liaison officer, editor)
Alexander Imig (treasurer, website editor)

ISSN: 2434-849X
DOI https://doi.org/10.37546/JALTSIG.CEFR

CEFR JOURNAL—RESEARCH AND PRACTICE
VOLUME 5

Table of Contents

Mission statement

The CEFR Journal is an online, open-access, peer-to-peer journal for practitioners and researchers. Our editorial advisory board 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.

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, and the Chinese Standard of English (CSE). 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 and direction. 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

Journal editorial team

Editing and proofreading team

Layout

CEFR JOURNAL—RESEARCH AND PRACTICE
VOLUME 5

Editorial

Fergus O'Dwyer

The open source movement emanating from software development promotes universal access, with resources made freely available for possible modification and redistribution. This spirit is found in the development of the CEFR, and the invitation in the article by Neus Figueras, David Little and Barry O'Sullivan to CEFR users to share their experience and projects that can be included in future iterations of the *Aligning Language Education with the CEFR: A Handbook*. The general outline of the handbook featured in this issue discusses the aim of unifying and aligning elements such as curriculum, teaching/learning materials, teaching approaches, teacher training, and assessment. This feeds into the multimedia orientation course on the CEFR Companion Volume (CV) by Monica Barsi and Teresa Betarello: language learning is viewed as an ongoing process where reception, production, interaction, and mediation blend into each other to make up linguistic competence. The action-oriented approach in part involves language learning events accumulating and building new knowledge to move forward. A promising element of the orientation course is the presentation of the CV in a way that illuminates the theoretical framework of the CEFR (2001) for younger audiences, who may have only accessed the 2020 publication.

The future development of the CEFR potentially involves disseminating best practice through collaboration between interested parties working in different contexts and the further expansion of professional networks. The article by Martina Hulešová outlines the positive impact the sign language project has had on the Czech deaf community, as they took part as consultants, validators, dissemination panelists so that they understand the use and usefulness of the project. What began as adaptation of proficiency scales morphed into a much more encompassing general *Framework of Reference for Sign Languages* and a specific *Reference Level Descriptors of Czech Sign Language.* These bilingual (written Czech and Czech Sign Language) resources facilitate coordination for teaching, syllabi, curricula, assessment, teaching and learning innovations. These can begin to overcome challenges presented by language modalities, the absence of standardisation, and critically the lack of research into, or recognition of Czech Sign Language.

The contribution by David Coniam, Michael Milanovic and Wen Zhao provides institutions of higher education guidance when considering the admission of Chinese students submitting results of the Chinese College English Test, which is based on the China Standards of English, and the *LanguageCert Test of English* (which, in turn, is based on the CEFR framework).

The final article discusses the introduction of a CEFR-aligned curriculum into the Israeli context. This required a change in teachers' mindsets alongside a long-term, multi-stage implementation plan. Tziona Levi and Simone Duval use the ripple effect metaphor to analyse professional learning communities, and specifically collaborative dialogues in learning the curriculum through social interaction, activating conceptual curriculum language as a mediational tool and 'languaging' the meaning making process. This mindset shift is in motion, with the call for experts to facilitate courses with teachers to prepare materials according to the guidelines of the CEFR-based curriculum.

We look forward to further conversation-starting articles in following issues of the *CEFR Journal*: please see the call for abstracts if interested in submitting.

— Dublin (Ireland), December 2022

CEFR JOURNAL—RESEARCH AND PRACTICE
VOLUME 5

Aligning Language Education with the CEFR: A Handbook

Neus Figueras, University of Barcelona

David Little, Trinity College Dublin

Barry O'Sullivan, British Council

https://doi.org/10.37546/JALTSIG.CEFR5-1

This article briefly presents the handbook Aligning language education with the CEFR, which was published in April 2022 and is freely available online. The publication of the CEFR Companion Volume with new descriptors (CEFR CV; Council of Europe, provisional version 2017, definitive version 2020) has caused quite a stir in the field of language education and prompted renewed interest in the content and applicability of the CEFR (Council of Europe 2001). Language assessment professionals in particular have discussed the many implications of the CEFR CV in different contexts and scenarios. The article explains why the handbook was developed and who it is for, describes the steps involved in aligning the different dimensions of language education with the CEFR, and explains how the handbook is organised.

Keywords: CEFR, CEFR CV, CEFR alignment, language assessment

1 Introduction

This article briefly presents the handbook mentioned in the title, which was published in April 2022, and freely available online. The publication of the CEFR Companion Volume (CEFR CV) with new descriptors (Council of Europe provisional version 2017, definitive version 2020) has caused quite a stir in the field of language education and prompted renewed interest in the content and applicability of the CEFR (Council of Europe 2001). Language assessment professionals in particular have discussed the many implications of the CEFR CV in different contexts and scenarios, which have been reported in this journal from its first issue.

The EALTA CEFR SIG held in Dublin in January 2018 (Little 2018) and the EALTA-UKALTA Symposium hosted by the British Council in London in February 2020 (O'Dwyer et al. 2020; Little and Figueras 2022) focused not only on the potential impact of the CEFR CV on language assessment but also on its implications for language education in general. Discussion at these events and in the various reports and publications mentioned suggested possible ways to increase transparency and collaboration in aligning different components of language education to the CEFR in different contexts and pointed to the need for a document which could support alignment with the greatly expanded descriptive scheme of the CEFR CV.

A small steering group was convened in early April 2020 to revisit the *Manual for Relating Language Examinations to the CEFR* (Council of Europe 2009) and to explore options for replacing it with a handbook that takes account of the recent publication of the CEFR CV, focuses not just on testing and assessment, and addresses a wide range of stakeholders. This steering group included key participants in the 2020 February symposium in London representing a range of organisations: Neus Figueras (EALTA), Barry O'Sullivan (British Council), Nick Saville (ALTE), Lynda Taylor (UKALTA), and David Little, with his extensive

knowledge of the Council of Europe.

Working with a wide range of colleagues across our respective organisations, the steering group has drafted a *Handbook for Aligning Language Education with the CEFR* that offers guidance and support for aligning curriculum guidelines, curricula, syllabuses, teaching materials and assessment with the CEFR.

The handbook was developed in three phases. First, the steering group agreed on its scope, drafted Chapters 1 and 2, an overview of the Council of Europe, the CEFR and the CEFR alignment process proposed in this handbook, and sent them to selected language education professionals for critical feedback. Next, specialists with extensive experience of working with the CEFR were invited to draft Chapters 3-6, which focus on the various stages of the CEFR alignment process. After that, the draft handbook was edited and circulated to the professionals who had provided feedback on the first drafts of Chapters 1 and 2; their comments and suggestions were incorporated in the final draft.

This article introduces the handbook, outlines its main purpose, describes its intended audiences, its contents and its structure, and concludes by encouraging prospective readers and users to share their experience and projects with the intention of incorporating them in a revised edition of the document. The February 2020 symposium concluded that the dissemination of best practice in aligning the elements of language education to the CEFR can promote collaboration between interested parties working in different contexts and contribute to further expansion of professional networks.

2 Why a handbook?

The February 2020 symposium concluded that alignment applies not only to language tests but to policy, curriculum guidelines, curricula, syllabuses, textbooks and other teaching/learning resources. Although these elements impact significantly on one another and on learning, curriculum developers, materials developers, teacher trainers and assessment specialists mostly work independently of one another. The need to consider these elements from a single unified perspective forms the basis of O'Sullivan's (2020) concept of the Comprehensive Learning System (CLS), which the CEFR (2001, 2020) also argues for in the inclusion in its title of the categories learning, teaching, assessment. O'Sullivan argues that the success of any learning system depends on the close alignment of elements that have traditionally been regarded as independent of one another: curriculum, teaching/learning materials, teaching approaches, teacher training, and assessment. Figure 1 below highlights the three core elements of the CLS:

1. **Curriculum** – informal as well as formal
2. **Delivery** – includes teacher selection, teacher training, accreditation, professional development and leadership; teaching and learning materials; the physical environment in which the delivery takes place
3. **Assessment** – includes developmental assessment (diagnostic, aspects of progress, formative, etc.) and judgemental assessment (placement, aspects of progress, achievement, proficiency, etc.)

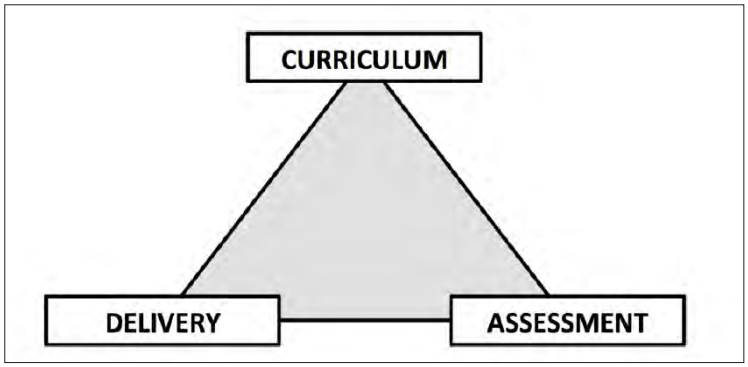


Figure 1. O'Sullivan's Comprehensive Learning System (CLS)

If one of these elements is in any way disconnected from the others, then the system is under threat.

The steering group used the CLS concept as the basis for developing and presenting the different activities of the handbook, which follows the overall structure of the *Manual for Relating Language Examinations to the CEFR* (published in its final version in 2009) but organises its contents so as to emphasise their relevance to different stakeholders working in different contexts.

2.1 Who is the handbook for?

The handbook aims to inform policy makers, teacher educators, teachers and other language education stakeholders and to support the more or less technical processes on which alignment depends. It has been prepared with two audiences in mind, each of which comprises a number of specialised subgroups, as shown below (the bulleted lists do not claim to be exhaustive). In both cases, the handbook is designed to help professionals working on their own, those working within an institution, and those with coordination responsibilities.

AUDIENCE 1	AUDIENCE 2
Those wishing to undertake **a practical or applied CEFR alignment exercise** in a particular context and for a particular purpose (e.g., to be able to make or evaluate a claim concerning CEFR alignment). This includes: • in-service teachers • materials developers • syllabus designers • pre-service teachers • test producers • textbook writers	Other stakeholders in education or in society at large who are **primarily concerned with policy matters and decision-making** relating to language education more generally (e.g., to evaluate the claims made by textbook or test publishers regarding CEFR level). This includes: • administrators • curriculum developers • education ministry personnel • employers • managers • policymakers • publishers • teacher trainers/educators

Table 1. ***Target audiences for the handbook***

3 What is involved in CEFR alignment?

Undertaking a CEFR alignment exercise involves one of two processes:

Process 1	OR	Process 2
Collecting evidence and developing an argument to show that an ***existing*** resource, e.g. a curriculum, a textbook or a language test, fulfils criteria derived from the CEFR.		Developing and documenting a ***new*** resource, e.g. a curriculum, a textbook or a language test, on the basis of the CEFR criteria.

Table 2. ***Alignment process types***

Process 1 and Process 2 both involve a series of well-established and largely sequential steps, or sets of procedures, as shown in Figure 2.

The first step in the alignment process is familiarisation, an essential stage at the outset of any alignment exercise. Experience from previous alignment studies has shown that familiarisation is a very useful, and sometimes necessary, preliminary activity for the other stages because it provides the opportunity to revise and reflect on specific elements relevant to the context of the alignment project.

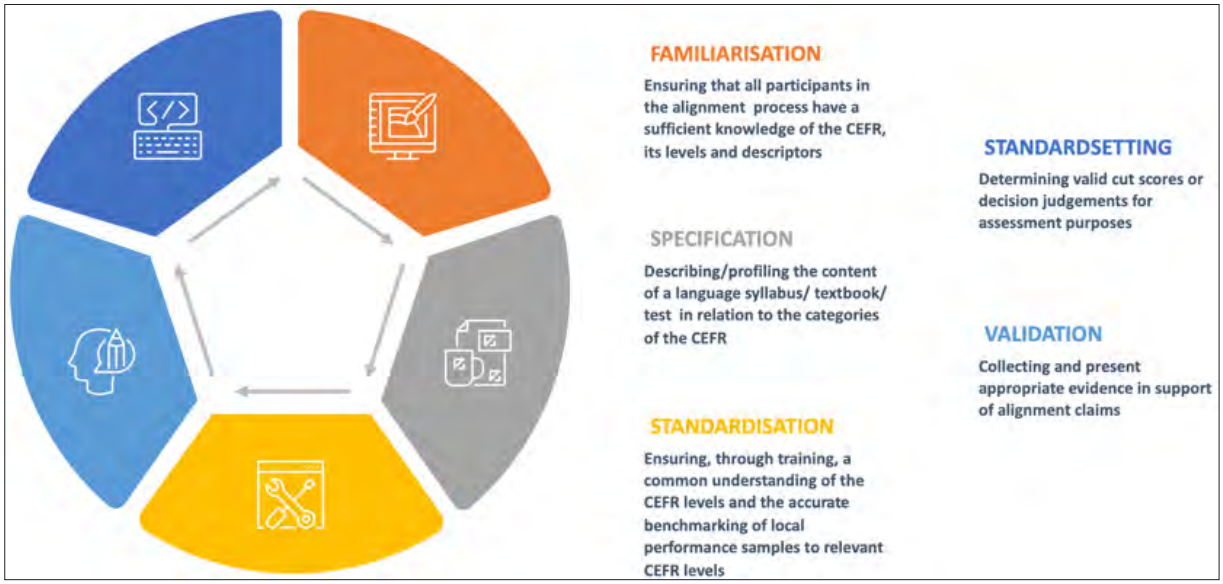


Figure 2. Steps in the alignment process

The next steps are **specification** and **standardisation**, which are required for CEFR alignment projects in most, if not all, contexts of use.

Specification implies the analysis of the content(s) of any resource, existing or new, in terms of approach and coverage in relation to the categories presented in the CEFR, whereas **standardisation** is the process of establishing that the main features of a given resource reflect a clear understanding of the relevant CEFR levels and descriptors.

For those involved in all aspects of establishing an empirically-based link between a curriculum, a set of materials (e.g. textbook or online course), or an assessment or test, **standard setting** procedures are a key requirement. For some of these contexts, the degree of standard setting activity required is likely to be less than we might expect for a test. The specific use of a test determines the choice and appropriateness of standard setting procedures.

Validation is best understood as the continuous process of quality monitoring in order to gather the evidence to support claims of CEFR alignment. Like familiarisation, validation is to some degree relevant to all the other steps in the alignment process – by demonstrating that all stages have been followed in an appropriate way, we establish evidence of the validity of subsequent claims of a link to the CEFR.

3.1 How is the handbook organised?

Each of the chapters focuses on one of the five procedures outlined above. Chapters 1 and 2 have been written to be accessible and directly relevant to both target audiences. Chapter 1 provides an essential introduction to the Council of Europe, the CEFR and the CEFR alignment process. Chapter 2 focuses on familiarisation as an important first step in any alignment project. Chapters 3-6 offer both general

and detailed guidance on the successive steps in an alignment process together with information on available tools and approaches to reporting on the activities undertaken.

The handbook refers to the CEFR 2001 and the CV 2020 as a *single* resource when undertaking the CEFR alignment process, and points readers and users to other publicly available resources which served as a basis for the handbook.

Each chapter begins with an explanation of the procedure and its importance, followed by general advice and practical activities that are relevant to all users regardless of language education context and their focus (e.g. curriculum, teaching, assessment). The remainder of each chapter offers more targeted guidance and practical activities specific to different contexts of use and their associated stakeholders, e.g. language teachers, curriculum designers, textbook writers, test producers. Readers and users can select from and focus on these according to need and context.

As the alignment process may be planned as a group approach and involve one or more coordinator(s) and participants, some practical advice and suggestions relating to these differing roles are included, including tasks to be completed with likely timings. At the end of each chapter, a "Notes for your own implementation" section provides a final reminder of the essential components of the activities presented.

The appendix contains photocopiable summary forms to use and complete. This additional practical tool can assist users in their ongoing monitoring and validation throughout the alignment process. The forms can be used as they are presented or adapted to fit the needs of a particular alignment approach or resource.

The handbook encourages reflection, and as users are the best judges of what is (and is not) possible or realistic within their specific context, it encourages them to tailor all activities to the specific context, taking account of available resources and limitations.

4 Future prospects

It is not possible to predict the success of the handbook presented here. Given the huge impact of the CEFR 2001 and the interest raised by the CEFR CV (COE 2020), we expect that this first edition will soon be put to use in alignment projects by many different stakeholders in the field of language education. We would like to invite those involved in such projects to share not only their outcomes, but also their views on the usefulness of the handbook. There are plans to host an event in early 2024 to present case studies and good practices in the use of the handbook so that the resulting suggestions and proposals can be incorporated in a future edition.

5 References

British Council, UKALTA, EALTA and ALTE. 2022. Aligning language education with the CEFR: A handbook. http://www.ealta.eu.org/documents/resources/CEFR%20alignment%20handbook.pdf (accessed 30 September 2022).

Council of Europe. 2001. Common European framework for languages: Learning, teaching, assessment. https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=0900001680459f97 (accessed 30 September 2022).

Council of Europe. 2009. Relating language examinations to the CEFR. A manual. Available at https://rm.coe.int/1680667a2d (accessed 30 September 2022).

Council of Europe. 2020. CEFR for languages: learning, teaching, assessment. Companion volume with new descriptors. https://rm.coe.int/common-european-framework-of-reference-for-languages-learning-teaching/16809ea0d4 (accessed 30 September 2022)

Little, David. 2018. The CEFR Companion Volume with new descriptors: Uses and implications for language testing and assessment. Report. VIth EALTA CEFR Sig. Trinity College Dublin. http://www.ealta.eu.org/events/Report%20on%20VIth%20EALTA%20CEFR%20SIG%20rev%2023.02.18.pdf (accessed 30 September 2022).

Little, David. and Neus Figueras (eds.). 2022. *Reflecting on the Common European Framework of Reference for Languages and its Companion Volume*. Bristol: Multilingual Matters

O'Dwyer, Fergus, Morten Hunke and Maria Gabriela Schmidt. 2020. The EALTA UKALTA 'Roadmap' conference – The CEFR: A road map for future research and development – meeting overview. https://cefrjapan.net/images/PDF/Newsletter/CEFRJournal-vol2.pdf#page=91 (accessed 30 September 2022).

O'Sullivan, Barry. 2020. The Comprehensive Learning System. London: British Council. Available at https://www.britishcouncil.org/sites/default/files/the_comprehensive_learning_system_new_layout.pdf (accessed 30 September 2022).

6 Biographies

Neus Figueras coordinated the development of foreign language curricula and certificate examinations for adult language learners in the Catalan ministry of education in Spain for over 20 years. She has been involved in a number of international research and development projects related to assessment at different education levels. She collaborates regularly with the Council of Europe in the uses and the dissemination of the CEFR. She is a founding member of EALTA (European Association for Language Testing and Assessment) and was the first president of the association. She is currently the coordinator of the EALTA CEFR SIG.

David Little is a Fellow Emeritus of Trinity College Dublin, Ireland. His principal research interests are the theory and practice of learner autonomy in second language education and the management of linguistic diversity in schools and classrooms. He has been a regular contributor to the Council of Europe's language education projects since the 1980s, especially in relation to the European Language Portfolio. In 2010, the National University of Ireland awarded him an honorary doctorate in recognition of his contribution to language education in Ireland and further afield.

Barry O'Sullivan is the Head of Assessment Research & Development at the British Council. He has undertaken research across many areas on language testing and assessment and its history and has worked on the development and refinement of the socio-cognitive model of test development and validation since 2000. He advises ministries and institutions on assessment policy and practice and is particularly interested in the communication of test validation and in test localisation. He is the founding president of the UK Association of Language Testing and Assessment (UKALTA) and holds honorary and visiting chairs at a number of universities globally.

CEFR JOURNAL—RESEARCH AND PRACTICE
VOLUME 5

A multimedia orientation course on the CEFR Companion Volume, going back to 2001, moving forward to 2020 and beyond

Monica Barsi, University of Milan

Teresa Bettarello, Provincial Centre of Adult Education, Milan

https://doi.org/10.37546/JALTSIG.CEFR5-2

In this contribution, we illustrate how we have designed our orientation course on the CEFR Companion volume: with what objectives (Part I) and what structure (Part II). The course is designed and realised in microlearning mode. Graphics, animation and approach to the subject matter make it enjoyable also for those who are approaching LS/L2 teaching or for self-learners who want to discover this document. Each of the 11 course units consists of a 5-10' video, equally usable from a mobile device, focusing on a particular aspect (all videos are visible on the website http://promoplurilinguismo.unimi.it and on youtube). All our references to the adoption and knowledge of the CEFR CV mainly concern our experience in Italy, although we have noticed many similarities with other countries both at school (at all levels) and outside school. 'Our' Italian society, just like 'our' European society, has in fact seen and continues to see the importance assumed by linguistic and cultural diversity growth of which it is necessary to become increasingly aware. The course is intended to be a tool for understanding the whole didactic idea of the CEFR CV, which envisages language levels from Pre-A1 to C2 but at the same time promotes education in plurilingualism, interculturalism, mediation and inclusiveness.

Keywords: Companion Volume, CEFR, glottodidactics, second languages, foreign languages, linguistic education, plurilingualism

1 Introduction

The publication of the Common European Framework of Reference for Languages Companion Volume (CEFR CV) 2020 marks a new milestone in the history of language education, comparable in scope to that of the first Common European Framework of Reference for Languages (CEFR) publication in 2001[1]. From the point of view of the reception of this document-instrument, the remarkable fact is that from 2020 onwards, it will not be as new as it was in 2001. This initial reflection accompanies any discussion of the *Companion Volume* (CV) that we would like to disseminate to teachers, scholars, and language

1. The bibliography on the CEFR 2001 is very extensive. Many contributions have been written in Italian, as well as in all other languages, so we refer to the Council of Europe (COE) website, which contains many articles, books, and dedicated websites; we think in particular of Beacco et al. 2016. Studies on the CV are at an early stage. We recall here the conference organised at the University of Milan in 2019, whose proceedings are published in *Italiano LinguaDue*: Barsi-Jardin 2020 and the contributions of Piccardo 2020, North 2020, Benedetti et al. 2020. The journal *Italiano LinguaDue* continues to publish contributions on the CV in Italy; see, for example, Cattaneo 2021 and Fratter 2021.

education professionals in the Italian version[2]. To do this, we have asked ourselves how to present the CV drawing on the theoretical framework of the CEFR (2001) without leaving out all those parts of the CEFR 2001 that are considered to have been acquired and are no longer included in it. In constructing an orientation course on the CV we have therefore had to and wanted to seek a balance to pass on to younger audiences a legacy of which they only see the latest publication of 2020. In this contribution, we would like to illustrate how we have designed our orientation course: with what objectives (Part I) and what structure (Part II)[3]. Of course, all our references to the adoption and knowledge of the CEFR mainly concern our experience in Italy, although we have noticed many similarities with other countries both at school (at all levels) and outside school.

2 Part I: Objectives of the course

The main purpose of the *Companion Volume* orientation course is to raise awareness of this new framework, as well as to disseminate it and use it. Because if the title, CV, indicates that this work cannot exist without the previous edition of 2001, which it nevertheless replaces while maintaining the characteristics of immediacy, transparency, coherence, dynamism and non-dogmaticity, to adapt to a changed and changing world, the CV significantly expands the vision of the CEFR 2001 by adapting to the needs arising from the increase in linguistic and cultural diversity within our societies.

To be clearer and to narrow down this analysis geographically, it is certainly clear to everyone that 'our world', meaning by 'our', the so-called Western part, the cradle of the CEFR, is constantly changing. Migrations lie at the origin of this state of affairs: for study, work, or family needs, to escape death from hunger-war-disease, people are constantly moving in or arriving in the West. The authors of the CV have revised the CEFR 2001 with a very realistic view. To give a few examples: the Pre-A1 level (with its descriptors) has been extensively introduced; it is a level that is necessary to teach the increasing number of students with little or no literacy in their mother tongue or who are linguistically very distant from the language they are learning. Special scales and descriptors have been drawn up: an entire chapter has been dedicated to sign language; for young learners, two age groups have been considered: 7-10 and 11-15. The real use of language has also led the authors of the CV to carefully consider online communication: equal but different from paper communication because it is written with a unique linguistic form (it is not only written, but it is also somehow spoken), it has become a scale apart in the communication activity Interaction. Finally, there are two major innovations that, in some way, alone challenge the hitherto conceived idea of teaching a second or a foreign language. We refer to mediation as a communication activity and strategy and to plurilingualism and pluriculturalism as a communicative competence. With them, the invitation is to look at language learning in a holistic way, which includes and involves behaviours and competences that are similar to Life skills (Fratter 2021) so much in demand in the labour market.

2. The Italian version is online in open access at *Italiano LinguaDue*: Council of Europe 2020b. The Italian translation is by Monica Barsi, Edoardo Lugarini and Anna Cardinaletti. The Italian version of the CEFR 2001 is referenced as Council of Europe 2002.
3. In the first investigation, we looked at online videos on various aspects of teaching with the CEFR. In particular, we looked at the sites https://www.lincdireproject.org/, especially the section on tutorials for the teacher (https://www.lincdireproject.org/lincdire-tutorial-videos/) and scenarios divided into different levels (https://lite.lincdireproject.org/it/all-scenarios-2/), and https://transformingfsl.ca for all available videos. We also relied on Piccardo et al. 2011, North 2014, North 2015, North et al. 2018, Mariani 2016, Piccardo 2014 and Piccardo 2019. The quality matrix was particularly useful in understanding how to build online materials on CEFR: 'A quality assurance matrix for CEFR use' (QualiMatrix CECR), European Center for Modern Languages (ECML) https://www.ecml.at/ECML-Programme/Programme2016-2019/QualityassuranceandimplementationoftheCEFR/tabid/1870/language/en-GB/Default.aspx

Taking all these new features into account, one can identify the objectives that can be achieved through the dissemination of the use of the CV. These can be divided into linguistic, socioeconomic, sociocultural and professional objectives.

The linguistic objectives aim to convey all the new features and updates contained in the CV. In addition, we felt it was essential to convey the way in which the CV replaces the traditional distinction that applied linguistics has made between the Chomskyan concepts of competence (hidden) and performance (visible) with proficiency, defined as the visible dimension of a person's competence resulting from a specific performance. Can Do statements indicate that practising the language is the best way to give evidence of being in the process of learning, to reach the mature stage of acquisition then.

The socioeconomic objectives aim to raise awareness of mediation and the related plurilingualism, both advocated by the CEFR as early as 2001 about plurilingual European citizens speaking more than two languages. The development of the areas of mediation and plurilingualism has actually had the effect of breaking down language barriers and creating more dynamic border regions, as well as international exchanges even between non-neighbouring countries[4]. The use of the CEFR 2001 applied to a single second or foreign language has made us forget the plurilingual purpose, which can be rediscovered through the potential of interlingual mediation. This type of mediation adds to the possibilities of communication beyond translation and interpreting[5]. In particular, we are keen to disseminate the idea of plurilingualism as a guarantee of the possibility for all speakers of different community languages to cooperate, for example, to carry out transnational professional training.

The sociocultural objectives aim to disseminate the CV to improve the communication and integration of all citizens within their own everyday environment, but above all beyond their national borders, with a view to European and non-European mobility. From a linguistic-legal perspective, the plurilingualism and consequent pluriculturalism promoted by the CV are also configured as vehicles for democratisation and the construction of a feeling of identification and belonging to Europe. Progression in linguistic level in each of the languages that a person knows begins with a reality of a few words ('I') and moves towards the complexity of the world that the 'I' includes and in which the 'I' is able to interact as a subject with rights and, therefore, to exercise active citizenship.

The professional objectives consist, in the Italian context, of disseminating the CV among the actors involved in the teaching of Italian as L2, that is, second and foreign language teachers and language facilitators in compulsory education as well as second and foreign language teachers in local authorities, companies, organisations linked to the labour market, associations and public schools; teachers at CPIAs (provincial centres for adult education); providers and certifiers of the Italian test for long-term resident immigrants; Italian teachers abroad, language assistants at schools, at Italian cultural institutes and foreign universities, and diplomatic representations, embassies and consular offices. As for foreign languages, the actors involved are the teachers of all languages as a second language in compulsory education, in private institutions, in universities and all students of any grade and any institution. The historical context seems favourable at this juncture: COVID-19 has forced almost all teachers to revise their way of teaching, a necessity that has led to an extraordinary intellectual opening of which we must absolutely take advantage.

There are many objectives that can be achieved by disseminating the use of the CV. However, there is a fundamental aim behind the course we designed: to disseminate the idea that the CV can be consulted in a targeted manner without having to be read in its entirety: like a recipe book, you can find useful information in it for planning a course or organising individual lessons. This aim lies at the origin of the structure of our course.

4. We recall here that, in addition to school education, mobility in vocational training has been a declared objective of the European Union since the early 2000s, as evidenced by the Kok Report, for instance (European Commission 2004).
5. On plurilingualism and mediation see Piccardo 2018, Piccardo-North-Goodier 2019, Piccardo-North 2020 and Crestani 2020. On mediation see also Coste-Cavalli 2019.

3 Part II: The structure of the course

The online orientation course on the *Companion Volume* was designed and delivered in microlearning mode on a Moodle platform[6]. Its graphics, animation and approach to the subject make the course enjoyable even for those who do not know the CEFR, or who are approaching second language teaching for the first time, or for autonomous learners who want to discover this document. Each of the 11 units of the course consists of a 5-10 minute video, which can be accessed from a mobile device: each video focuses on a particular aspect and is followed by a test of three self-assessment questions. At the end of the course, which lasts approximately three hours and can be carried out over an indefinite period, users will answer a final set of questions. The successful completion of this final test, in the context of Lifelong Learning courses, will provide users with a training certificate.

The table of contents gives an overview of the course we have designed:

1. The CV as a user-friendly tool
    1.1. The guiding colours
    1.2. Common Reference Level descriptors
2. The CEFR descriptive scheme
    2.1. An indispensable overview
    2.2. Communication as a rainbow
3. Changes in 2020
    3.1. General presentation
    3.2. Presentation of appendices
4. The Inclusive Model of the CV

As shown above, the first two stages concern the system of descriptors for all scales of communicative competencies, activities, and strategies, as it appeared in 2020. By contrast, the final two stages deal with innovations and changes in the CV compared to the CEFR 2001.

4 The first stage

In the first stage, entitled The *Companion Volume as a user-friendly tool*, we focused on the structure and graphics of the CV, showing how the colours are a guide to orientate the reader. In the online course, we have therefore maintained, where possible, this colourful and very intuitive distinction between the parts to make the user familiar with them. The theoretical framework of the CEFR remains unchanged, and we refer to the description of the common framework in 2001[7]. We have also provided an explanatory description of the table of contents of the CV, which is formulated as follows:

- The introduction presents the CV and the changes that have taken place since 2001.

6. Concept by Monica Barsi and Teresa Bettarello and multimedia realisation by Angelisa Leonesio (InterARTactivity). The course on Moodle is available at the University of Milan and ANILS and is accessible only to registered students. The videos have been made available in open access on http://promoplurilinguismo.unimi.it and YouTube. The course has been awarded the 2020 European Language Label.
7. We have quoted this passage as a summary: "Language use, embracing language learning, comprises the actions performed by persons who as individuals and as social agents develop a range of competencies, both **general** and in particular **communicative language** competencies. They draw on the competencies at their disposal in various **contexts** under various **conditions** and under various **constraints** to engage in **language activities** involving **language processes** to produce and/or receive **texts** in relation to **themes** in specific **domains**, activating those **strategies** which seem most appropriate for carrying out the **tasks** to be accomplished. Monitoring these actions by the participants leads to the reinforcement or modification of their competencies" (CEFR 2001: 9).

- The four chapters that follow are devoted to communicative language activities and their strategies: reception, production, interaction, and mediation with the relative scales.
- Another chapter is dedicated to plurilingual and pluricultural competence.
- A penultimate chapter is dedicated to linguistic, sociolinguistic, and pragmatic competence and another chapter on linguistic, sociolinguistic, and pragmatic competence in sign language.
- Lastly, there are extensive appendices with expansions on some topics and where the updated self-assessment grid is located.

To help the reader refer to the CV in a targeted way, from its general structure, we have then focused on the cornerstone of its internal structure, which is the scales and descriptors. Each scale of descriptors is preceded by a summary of the content covered by the descriptors themselves (the so-called 'operationalised key concepts'), which is a very useful element for understanding in brief what a scale contains without having to read all the descriptors, making the selection of the scale or scales by the user much faster and more effective. The example given below is the scale for the communicative language activity 'Watching TV, films and videos' (Council of Europe 2020a: 52-53), for which the following explanation has been given:

Table 1. ***Rationale of the scale 'Watching TV, films and videos' (Council of Europe 2020a: 52)***

This scale includes live and recorded video material plus, at higher levels, films. Key concepts operationalised in the scale include the following: • following changes of topic and identifying main points; • identifying details, nuances and implied meaning (C levels); • delivery: from slow, clear standard usage to the ability to handle slang and idiomatic usage.

Regarding descriptors, new ones have been added to those of 2001. We have specified that the previous descriptors are not distinguished from the ones newly added (as was the case in the 2018 prototype) and that their extension takes place at all levels. It was then considered important to bring to mind what we can call the 'method of descriptors' because it was observed that, in many teaching situations, the full potential of the Can Do statement is not exploited. Thus, some basic elements have been recalled:

- descriptors contain mainly concrete and observable aspects of communicative performance;
- scales and descriptors serve firstly to identify which scales are relevant to meet the needs of a particular group of learners and only secondly to establish what level these learners need to achieve to reach their goals;
- descriptors' immediacy and transparency, in the sense of clarity and definition, allow the learners themselves, if they are adults, to negotiate with the teacher what they need to become able to use the language, rather than just to study it for no real purpose.

It was important to point out that descriptors are examples of typical language use and that, since the CEFR 2001, they are called 'illustrative' for this reason. They provide a basis for the development of behaviours that are appropriate to a given context but are not prescriptive in themselves. They are a basis for reflection, discussion and planning based on needs analysis. It has been pointed out that it is important to work with descriptors, even by modifying them, because they provide a common metalanguage that facilitates networking, the creation of communities of practice by groups of teachers and, only lastly, the definition of transparent criteria for evaluation.

One of the key elements we have returned to concerns programming with descriptors. In the classroom, the teacher derives the task that is to be proposed to the students from the descriptor – or

from several descriptors combined together – and, more generally, the teacher derives the programme of a whole course from quite a large number of selected descriptors.

At this point the structure of the descriptor is shown by listing the following parts:

- The communicative language activity involved;
- The textual type/form, which may sometimes be appropriate for several language levels;
- The subject matter of the communication: what is being talked about?;
- The conditions and constraints under which the learner has to learn to operate. The same task, in fact, may be required at different levels but under different conditions, i.e., more or less favourable or more or less complex (e.g., leaving time to think before speaking, speaking with a collaborating interlocutor, in neutral circumstances or challenging situations). Such conditions and constraints, of course, increase as the level rises.

As an example, we have illustrated the descriptor for level B1 and level C2 in the scale 'Understanding as a member of a live audience' (Council of Europe 2020a: 50-51):

Table 2. Descriptors B1 and C2 of the scale 'Understanding as a member of a live audience' (Council of Europe 2020a: 50)

B1	Can follow in outline straightforward short talks on familiar topics, provided these are delivered in clearly articulated standard language or a familiar variety.

C2	Can follow specialised lectures and presentations employing colloquialism, regional usage or unfamiliar terminology.

In both descriptors, we recognise the four categories that have been listed above and which we break down here in the following table:

Table 3. Breakdown of descriptors B1 and C2 of the scale 'Understanding as a member of a live audience' (Council of Europe 2020a: 50)

	B1	C2
Activity	can follow in outline	can follow
Textual type/form	straightforward short talks	lectures and presentations
Subject matter	on familiar topics	specialised
Conditions and constraints	provided these are delivered in clearly articulated standard language or a familiar variety.	employing colloquialism, regional usage or unfamiliar terminology.

It seemed fundamental at this point to remember the distinction between the descriptors of the communicative language activities that indicate *what* a person is doing at a linguistic level and the descriptors of the communicative language competencies that indicate *how* s/he is doing it. The function of the descriptors for the communicative language strategies for the four modes of communication (reception, production, interaction, mediation) is, on the other hand, a connection between the two groups of descriptors because they link the *what* to the *how* on the path to proficiency.

Proficiency is a key concept that we explored. The question we addressed was, 'How do we progress?' In summary, we answered 'by practising', as the action-oriented approach claims. In the beginning, the learners know little or nothing. Their 'bridge' to the second or foreign language is them with their knowledge; but then, by practising the language, the learners learn, develop new skills, acquire new strategies and, therefore, increase their proficiency. In other words, the learners progress because they

practise the language in performing tasks and also through reflection with the Can Do statements. In this process, the learners also use their general skills, their communicative language skills and their plurilingual and pluricultural competence. We have highlighted how this learning process is circular, self-sustaining and self-stimulating from the lowest levels according to the principle that a person is always able to communicate something. In this regard, we felt it was important to point out that this principle must be clear to both parties involved in the learning process, teachers and students so that this document serves to learn and teach before assessing, just as the subtitle of the CEFR 2001, and now the CV of 2020 indicates.

Regarding assessment, we wanted to insist on using Can Do statements as a reference that teachers and students can share to measure progress. The descriptors facilitate the work of both parties: having established the goal to be reached, the teacher can plan 'backwards' what to do (backplanning). Once the goal has been achieved, the learner can become aware of what s/he can do through the descriptors and a 'backward' self-assessment. To illustrate how progress is made, we have shown the (Council of Europe 2020a: 36). In the concentric circles, the levels and their descriptors expand and contain the previous ones.

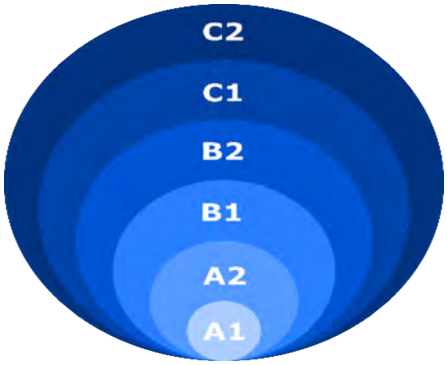


Figure 1/**CV Figure 3**. Representation of the CEFR Common Reference Levels (COE 2020a: 36).

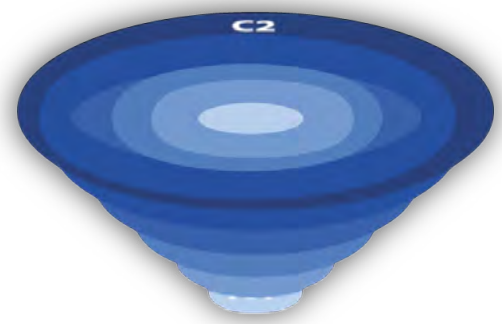


Figure 2. Rendering the representation in a slightly different manner, attempting to mirror the development required by the learner.

5 The second stage

At the second stage, entitled The CEFR descriptive scheme, we focused on the following figures within the original document of the *Companion Volume*, as follows: CV Figure 1 (Council of Europe 2020a: 32), CV Figure 2 (Id.: 34), CV Figure 4 (Id.: 36) and CV Figure 5 (Id.: 36) within the original CV, which dynamically represent the four modes of communication and the complexity of proficiency. We have recalled how already in 2001, the four skills of listening and reading/speaking and writing, i.e., oral and written reception and oral and written production, had been replaced and included in the current 'four modes of communication', i.e., reception, production, interaction and mediation. From the four skills to the four modes of communication, the change was fundamental to overcome the linear view of learning, based on the simple sum of parts that, it was assumed, could be broken down and taught and assessed separately. In the CEFR, the vision of didactics has always been completely holistic: the whole does not correspond to the sum of its parts; it is a unicum; the learners are 'social agents', who mobilise in the relationship with other individuals, their linguistic resources – not only formal ones – to communicate, that is to negotiate and co-construct the sense of the exchange. The movement of the arrows is explained in CV Figure 2 (Council of Europe 2020a: 34): interaction includes both reception and production, but is more than the sum of these two parts; mediation is even all-encompassing: it includes reception and production and, frequently, interaction.

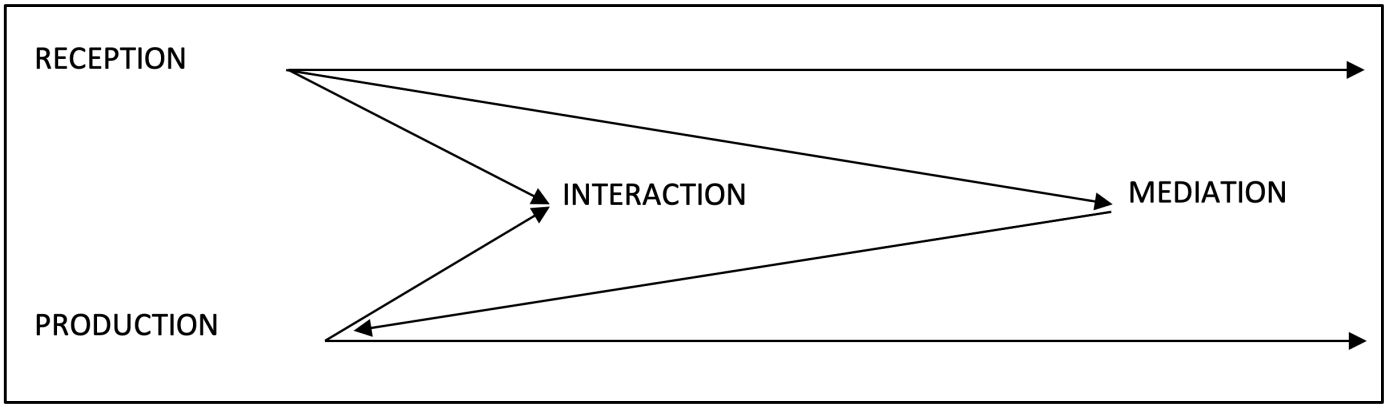


Figure 3/**CV Figure 2**. Representation of the relationship between reception, production, interaction, and mediation (Council of Europe 2020a: 34)

At this point we are taking a closer look at the descriptive scheme, i.e., CV Figure 1 (Council of Europe 2020a: 32):

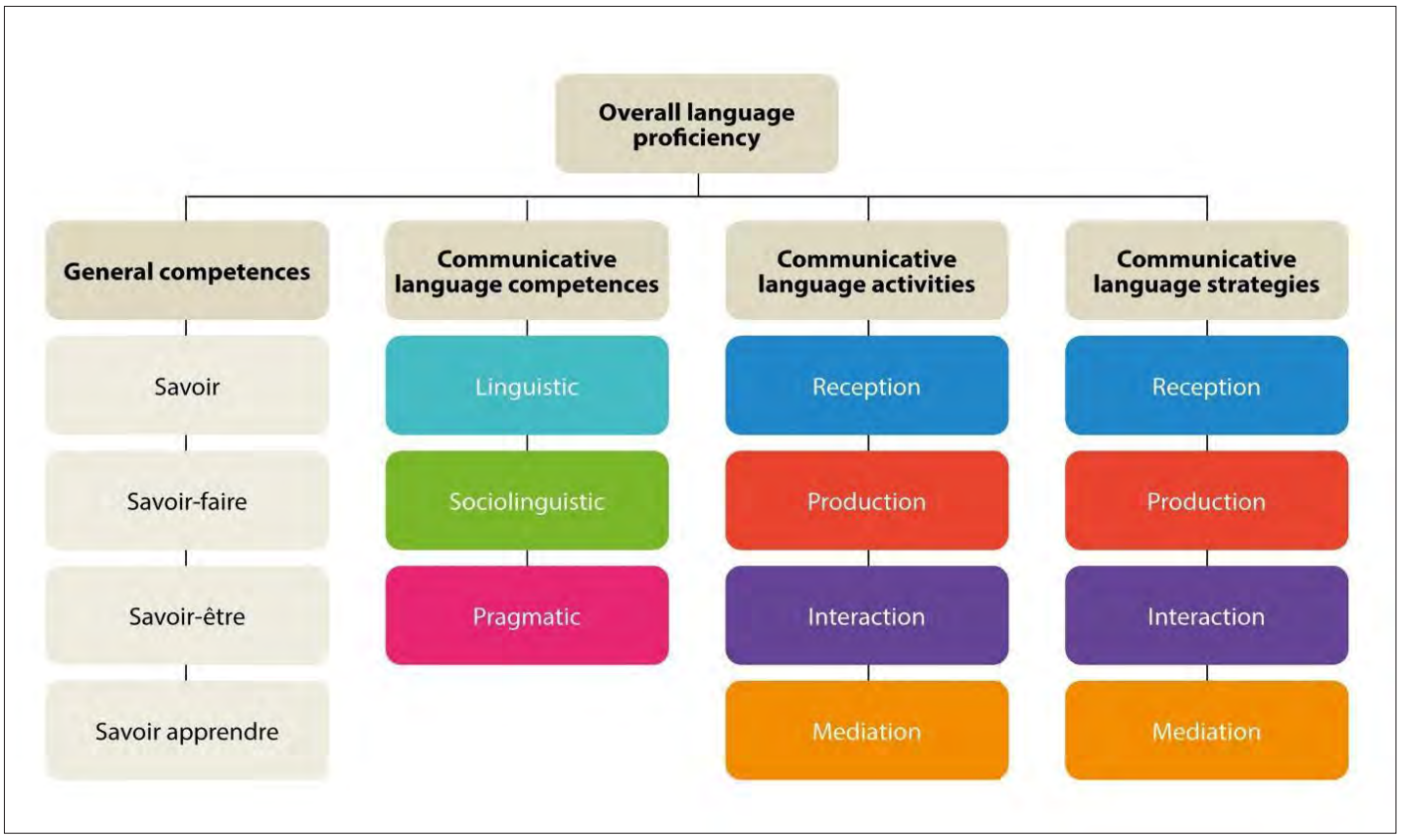


Figure 4/**CV Figure 1**. Representation of the CEFR descriptive scheme (Council of Europe 2020a: 32)

The explanation of the descriptive scheme allows us to recall the theoretical framework of the CEFR 2001. The learners/speakers are social agents who 'produce messages' (written, oral or combined and/or in sign language) through communicative language activities and strategies, mobilised by their personal general and communicative language competencies. In short, the learners/speakers communicate by bringing their whole self into play, not only the notions they learn in class: for example, they use their previous knowledge, their being in the world, making comparisons with their culture of origin and the

culture of the language being studied. This is why we have warned that, at the same level of achievement, two students in a B1 class will have arrived there through very different dynamics, such as according to their own personality, background and expectations.

From this perspective, language is no longer just the product of learning but an ongoing process. In this process, the ways of communicating are the expression of both skills and activities, which are linked by the strategies that are, to all intents and purposes, a connecting element. The question we formulated was, 'What are the best didactics to use'? According to the CEFR, it is an action-oriented one: the learners mobilise all their competencies to effectively complete tasks involving communication activities in collaboration with other people. In this way, they do not accumulate knowledge but practise it, even when it is very limited and, by practising, they build new knowledge and thus move forward, progressing in the different competences, often summarised as 'proficiency', that is the concept we highlighted in the first stage.

To explain the vision of the CV in its deepest dimension, we have chosen to use the image of the rainbow in CV Figure 4 and CV Figure 5 (Council of Europe 2020a: 36), which metaphorically represent general language competence and its holistic vision of action-oriented learning: if in the rainbow, the colours are clearly distinguishable but the boundaries are poorly marked – they 'blend' into each other, creating a coloured continuum-in the same way, in the descriptive scheme of the CV, general linguistic competence is made up of pieces of different colours which, in a dynamic relationship, contribute to the 'language process'.

Figure 5/**CV Figure 4**. Representation of a rainbow (Council of Europe 2020a: 36)

Figure 6/**CV Figure 5**. Representation of the conventional six colours (Council of Europe 2020a: 36,)

Each of these pieces has descriptor scales, i.e., contributes in its own way to the language process, but at the same time, all of them together must be considered a continuum.

The link with the CEFR 2001 for the general competences seemed to us to be necessary because of the perspective given by colours in the CV. In fact, the savoir, 'savoir être', 'savoir faire', 'savoir apprendre' described in the CEFR 2001 are not coloured and still have no descriptors in 2020. Therefore, the explanations of the CEFR 2001 have been summarised with a visualisation of the four types of knowledge.

6 The third stage

The third stage listed the 2020 changes, which we commented on in this order:

1. Pre-A1 level.
2. Modified descriptors (Appendix 7).
3. New detail at 'plus' levels (It is highlighted that the transition to the plus level is indicated by a white horizontal line).
4. Phonological proficiency (The concept of intelligibility is explained).

5. Mediation (A brief explanation of mediation is given, with reference to a specific section currently under development).
6. Plurilingual and pluricultural competence (A reference is also made to the CARAP[8]).
7. Creative texts (Three new illustrative scales are listed: 'Reading as a leisure activity;' 'Expressing a personal response to creative texts (including literature)'; 'Analysis and criticism of creative texts (including literature).'
8. Online interaction[9].
9. Sign language (It is explained that all descriptors have been made usable in sign language and 14 specific scales for sign language skills have been created).
10. Young learners (Reference is made to the section on the Council of Europe website: Council of Europe 2018a, 2018b).

The appendices have been carefully presented so that they are not overlooked, given their position in the *Companion Volume*. In particular, Appendices 2, 3, 4 and 5 have been described. This part of the Volume is, in fact, particularly rich. It contains, for example, the self-assessment grid, which, as well known, has been used in many cases as the only tool taken from the CEFR 2001.

7 The fourth stage

In the fourth stage, we felt it was essential to highlight the educational idea of the *Companion Volume*, starting with one of the first paragraphs:

> The Council of Europe hopes that the development in this publication of areas, such as mediation, plurilingual/pluricultural competence and signing competences will contribute to quality inclusive education for all and the promotion of plurilingualism and pluriculturalism. (Council of Europe 2020a: 11)

'Quality inclusive education for all' is the starting and ending point for teachers and learners, a virtuous circle to be covered in and out of the classroom. All of the changes seen in the different sections dedicated to them contribute to the inclusive model of the CV: sign language, the abandonment of the figure of the ideal native speaker, the Pre-A1 level for all, including semi-illiterates, the focus on children and pre-adolescents, and the importance given to mediation: all these aspects highlight the CV's educational idea that it is always possible to communicate and therefore to include people.

Plurilingual competence is also explored. In particular, we focused on plurilingual profiles, where the imbalance in skills is a value. This is because the learners/speakers who study/know more than one language are aware of their imbalances. Moreover, the learners/speakers are more or less consciously helped by what they can already do in (other) languages to strive for balance in the competencies of each language or to understand their primary needs in using the language they are learning. We have explained how the CV has fully embraced a holistic and dynamic view of competence in different languages. In fact, the underlying idea is not to separate the languages we know because they are all by nature always present in the speaker's mind, even if not at the same level. Precisely for this reason, they come into play in the learning of other languages through similarities and divergences.

8. IL CARAP, Un quadro di riferimento per gli approcci plurali alle lingue e alle culture. Competenze e risorse (https://riviste.unimi.it/index.php/promoitals/article/view/2823/3026), available in its Italian translation in Italiano LinguaDue. Italian version of *FREPA – Competences and resources* CARAP/FREPA - A Framework of Reference for Pluralistic Approaches.
9. On online interaction, see Langé et al. 2020. The authors of the book have conducted a large-scale trial of the descriptors throughout Italy.

8 Conclusion

Translated over time into 40 languages, the CEFR 2001 has established itself in Europe as a soft law in the field of certification. This has created an ever-growing echo that has crossed European borders thanks to the highly effective sharing of its system for assessment and self-assessment. In fact, the structure of the CEFR 2001 clearly suggested the relationship between the descriptor scales and assessment, to which Chapter 9 was dedicated. It probably described less clearly the relationship between the descriptor scales and the action-oriented approach, i.e., teaching through tasks. Much space was also devoted to the task, but the concept was new, and the relationship between descriptor/descriptors and the task was not exemplified.

As Piccardo and North explain in their 2019 book (Piccardo-North 2019), the action-oriented approach, described very briefly in the CEFR 2001, has been an area of research for almost two decades[10]. Teachers who have adopted the action-oriented approach have enriched its conceptual framework by reproducing real-life situations in their classrooms to make students use language with a purpose, i.e., in a way that is immediately meaningful and in which the reason is not postponed to the future. They have identified and assigned tasks to their learners, prompting them to mobilise their resources and skills and, consequently, develop essentially but not exclusively linguistic competences, as well as appropriate learning skills. They have used the CEFR descriptors to construct the tasks and the teaching sequences in which the tasks have been embedded. They have innovated assessment by basing it on descriptors. They have shared with their classes the need to recall all the knowledge they had acquired to immerse themselves in this methodological approach. This approach was not only an experiment, but above all, a paradigm shift that had already been initiated by the communicative turn away from the previous foreign and second language learning-teaching methods.

The online orientation course presented in this article fully embraces this paradigm shift and a perspective determined by the complexity of reality to transform the needs of individuals and groups into glottodidactic action in the context of 'our' societies where linguistic and cultural identity assertion is growing. In fact, any kind of difference can be safeguarded and enhanced through the inclusive model promoted by the *Companion Volume* and its premise, namely that inclusiveness always goes along with dignity, which is a concept that represents – in every intellectual process – a crucial motivation element.

9 References

Barsi, Monica & Anne Jardin (eds.). 2020. Verso nuove frontiere della comunicazione: l'interazione on line e la mediazione. Riflessioni sul *Companion Volume* del QCER e il Progetto Erasmus+ DELCYME [Towards new frontiers of communication: Online interaction and mediation. Reflections on the *Companion Volume* of the CEFR and the Erasmus+ DELCYME Project], *ItalianoLinguaDue* 12 (1). https://riviste.unimi.it/index.php/promoitals/article/view/13942 (accessed 1 December 2022).

Beacco, Jean-Claude, Michael Byram, Marisa Cavalli, Daniel Coste, Mirjam Egli Cuenat, Francis Goullier & Johanna Panthier. 2016. Guida per lo sviluppo e l'attuazione di curricoli per una educazione plurilingue e interculturale [Guide for the development and implementation of curricula for plurilingual and intercultural education]. Consiglio d'Europa, Strasburgo, Italian version: *Italiano LinguaDue* 8 (2). https://riviste.unimi.it/index.php/promoitals/article/view/8261/7882 (accessed 1 December 2022).

Benedetti, Fausto, Letizia Cinganotto & Gisella Langé. 2020. *L'interazione online nel Companion Volume del Quadro Comune europeo di Riferimento per le lingue: un progetto pilota italiano* [*Online interaction in the Companion Volume of the Common European Framework of Reference for Languages: An Italian pilot project*]. Firenze: INDIRE. http://www.indire.it/wp-content/uploads/2020/07/Volume-Companion-07.07.pdf (accessed 1 December 2022).

10. See Bourguignon 2010 and Hoerath 2020 to cite just two examples in the French-speaking world.

Bourguignon, Claire. 2010. *Pour Enseigner les Langues avec le CECRL, clés et conseils* [*Teaching Languages with the CEFR, keys and tips*]. Paris: Delagrave.

Cattaneo, Antonella. 2021. La competenza plurilingue e pluriculturale alla Scuola Europea di Varese. Un esempio di applicazione dei nuovi descrittori del *Volume complementare* del QCER [Plurilingual and pluricultural competence at the European School in Varese. An example of the application of the new descriptors of the CEFR *Companion Volume*], *Italiano LinguaDue* 13 (1). 975-1020. https://riviste.unimi.it/index.php/promoitals/article/view/15923/14260 (accessed 1 December 2022).

Council of Europe. 2001. *Common European Framework of Reference for Languages: Learning, teaching, assessment* (CEFR). English version: Cambridge University Press. https://rm.coe.int/16802fc1bf (accessed 1 December 2022).

Council of Europe. 2002. *Quadro comune europeo di riferimento per le lingue: apprendimento, insegnamento, valutazione* (QCER) [*Common European Framework of Reference for Languages: learning, teaching, assessment* (CEFR)]. Versione italiana: RCS Scuola Spa, La Nuova Italia, Oxford University Press.

Council of Europe. 2018a. *Collated Representative Samples of Language Competences Developed for Young Learners aged 7-10 years*, Strasbourg: Council of Europe. https://rm.coe.int/16808b1688 (accessed 1 December 2022).

Council of Europe. 2018b. *Collated Representative Samples of Language Competences Developed for Young Learners aged 11-15 years*, Strasbourg: Council of Europe. https://rm.coe.int/16808b1689 (accessed 1 December 2022).

Council of Europe. 2020a. *Common European Framework of Reference for Languages: learning, teaching, assessment – Companion Volume.* English version: Strasbourg: Council of Europe. https://rm.coe.int/common-european-framework-of-reference-for-languages-learning-teaching/16809ea0d4 (accessed 1 December 2022).

Council of Europe. 2020b. Quadro comune europeo di riferimento per le lingue: apprendimento, insegnamento, valutazione. Volume complementare [*Common European Framework of Reference for Languages: learning, teaching, assessment – Companion Volume*], *Italiano LinguaDue*. https://riviste.unimi.it/index.php/promoitals/article/view/15120/13999 (accessed 1 December 2022).

Coste, Daniel & Marisa Cavalli. 2019. Educazione, mobilità, alterità. Le funzioni di mediazione della scuola [Education, mobilité, altérité. Les fonctions de médiation de l'école], Italian version: Silvia Scaramella, *Italiano LinguaDue* 11 (1). https://riviste.unimi.it/index.php/promoitals/article/view/11877 (accessed 1 December 2022).

Crestani, Valentina. 2020. Mediare in 'Leichte Sprache' in tedesco e in italiano [Mediating in 'Leichte Sprache' in German and Italian], *Italiano LinguaDue* 12 (1). 586-602. https://riviste.unimi.it/index.php/promoitals/article/view/13947/13089 (accessed 1 December 2022).

European Commission. 2004. *Relever le défi : la stratégie de Lisbonne pour la croissance et l'emploi : rapport du groupe de haut niveau présidé par M. Wim Kok [Facing the Challenge: The Lisbon strategy for growth and employment. Report from the High Level Group chaired by Wim Kok November 2004]* Brussels: European Commission Publications Office. https://op.europa.eu/fr/publication-detail/-/publication/88b6bc81-e3ad-4156-960f-f549369aa9d4 (accessed 3 December 2022).

Fratter, Ivana. 2021. Le Life skills nel *Volume complementare* del QCER [Life skills in the CEFR Companion Volume], *Italiano LinguaDue* 13 (1). 157-168. https://riviste.unimi.it/index.php/promoitals/article/view/15863/14223 (accessed 1 December 2022).

Hoerath, Elisabeth. 2019. *La grammaire en action. Enseignement efficace de la grammaire dans le cadre d'une perspective actionnelle.* Français langue seconde (FLS) [*Grammar in action. Effective teaching of grammar from an action perspective. French as a second language* (FSL)]. Ontario, Canada. https://transformingfsl.ca/wp-content/uploads/2019/07/FSL_GrammaireenAction-2.pdf (accessed 1 December 2022).

Langé, Gisella, Letizia Cinganotto & Fausto Benedetti. 2020. Interazione online: una sperimentazione italiana [Online interaction: an Italian trial]. *Italiano LinguaDue* 12 (1). 603-612. https://riviste.unimi.it/index.php/promoitals/article/view/13948 (accessed 1 December 2022).

Mariani, Luciano. 2016. (Ri)Visitare l'approccio teorico-metodologico del Quadro Comune Europeo: materiali di formazione per insegnanti [(Re)exploring the theoretical-methodological approach of the Common European Framework: training materials for teachers]. *Italiano LinguaDue* 8 (2). 281-306. http://riviste.unimi.it/index.php/promoitals/article/view/8188 (accessed 1 December 2022).

North, Brian. 2014. *The CEFR in practice*. Cambridge: Cambridge University Press.

North, Brian (ed.). 2015. *Inventaire des contenus clés des niveaux du CECRL* [*Inventory of key contents of the CEFR levels*]. London: Eaquals.

North Brian. 2020. The CEFR renewed: Inspiring the future of Language education. *Italiano LinguaDue* 12 (1). 549-560. https://riviste.unimi.it/index.php/promoitals/article/view/13945/13087 (accessed 1 December 2022).

North, Brian, Mila Angelova & Elzbieta Jarosz. 2018. *Language Course Planning*. Oxford: Oxford University Press.

Piccardo, Enrica. 2014a. *Du communicatif à l'actionnel : un cheminement de recherche* [*From Communicative to Action-oriented: a Research Pathway*]. Version française. Service des programmes d'études Canada. https://transformingfsl.ca/fr/components/du-communicatif-a-lactionnel-un-cheminement-de-recherche (accessed 1 December 2022).

Piccardo, Enrica. 2014b. *From Communicative to Action-oriented: A Research Pathway*. English version. Curriculum Services Canada. https://transformingfsl.ca/wp-content/uploads/2015/12/TAGGED_DOCUMENT_CSC605_Research_Guide_English_01.pdf (accessed 5 December 2022).

Piccardo, Enrica. 2018. *Mediation "A paradigm shift in language education"*. Strasbourg: Council of Europe. https://rm.coe.int/mediation-a-paradigm-shift-in-language-education-piccardo/16808ae720 (accessed 1 December 2022).

Piccardo, Enrica. 2019. TIRF language education in review – The Common European Framework of Reference (CEFR) in language education: Past, present, and future. Monterey, CA & Baltimore, MD: TIRF & Laureate International Universities. https://www.tirfonline.org/wp-content/uploads/2020/03/TIRF_LEiR_CEFR_Final.pdf (accessed 1 December 2022).

Piccardo, Enrica. 2020. La mediazione al cuore dell'apprendimento linguistico per una didattica 3.0 [Mediation at the heart of language learning for didactics 3.0], *Italiano LinguaDue* 12 (1). 561-585. https://riviste.unimi.it/index.php/promoitals/article/view/13946/13088 (accessed 1 December 2022).

Piccardo, Enrica, Marie Berchoud, Tiziana Cignatta, Olivier Mentz & Malgorzata Pamula. 2011. *Pathways through Assessing, Learning and Teaching in the CEFR*. Strasbourg: Council of Europe.

Piccardo, Enrica & Brian North. 2019. *The Action-oriented approach. A Dynamic Vision of Language Education*. Bristol: Multilingual Matters.

Piccardo, Enrica & Brian North. 2020. The Dynamic Nature of Plurilingualism: Creating and Validating CEFR Descriptors for Mediation, Plurilingualism and Pluricultural Competence. In Lau, Sunny Man Chau & Saskia van Viegen (eds.), *Plurilingual Pedagogies*. *Educational Linguistics* (vol. 42), 279-301. Cham: Springer. https://doi.org/10.1007/978-3-030-36983-5_13 (accessed 1 December 2022).

Piccardo, Enrica, Brian North & Timothy Goodier. 2019. *Broadening the scope of language education: mediation, plurilingualism, and collaborative learning: the CEFR Companion Volume*, Journal of e-Learning and Knowledge Society. https://www.je-lks.org/ojs/index.php/Je-LKS_EN/article/view/1612/1032 (accessed 1 December 2022).

10 Biographies

Monica Barsi is a full professor of French and glottodidactics as well as head of the Study Programme – Academic Board for the Bachelor's Degree Programme in Foreign Languages and Literatures and for the Master's Degree Programme in Extra-European Languages and Literatures – at the Università degli Studi di Milano (University of Milan). In 2020, she co-translated (with Edoardo Lugarini) the *Companion Volume* into Italian. She has received the European Language Label for the online course on the CV (http://promoplurilinguismo.unimi.it). Her research and publications include the relations between Italy and France in the Renaissance, the history of school publishing, the bilingual French-Italian lexicography.

Teresa Bettarello is a freelance teacher of Italian as a second language, currently working with Centro Provinciale di Istruzione agli Adulti Milano (Provincial Centre of Adult Education, Milan) a public organisation that provides Italian language courses to foreigners. After a life as a journalist, spent covering novels and general culture, she approached teaching Italian to foreign students. She obtained the CEDILS certification in 'Italian language teaching to non-native speakers' (Ca' Foscari, Venice), to then achieve the Master Promoitals – 'Promotion and Teaching of the Italian language and culture to non native speakers' (Università degli Studi di Milano).

CEFR JOURNAL—RESEARCH AND PRACTICE
VOLUME 5

CEFR and CSE comparability study: An exploration using the Chinese College English Test and the LanguageCert Test of English

David Coniam, LanguageCert
Michael Milanovic, LanguageCert
Wen Zhao, Jinan University, Guangzhou

https://doi.org/10.37546/JALTSIG.CEFR5-3

This paper outlines how different studies can be brought together to reveal how two separate examinations, based on different assessment frameworks, may be compared. The paper reports on data obtained from a cohort of a comparatively large sample of Chinese university test-takers who took two separate tests – the Chinese College English Test (CET), which is linked to the descriptive scales of the CSE (China Standards of English) and the LanguageCert Test of English (LTE) linked to the descriptive assessment framework of the CEFR (Common European Framework of Reference for Languages). In addition to the test-taking, two further studies were conducted.

Analysis indicated that it was possible to make reasonably definitive pronouncements about the comparability of the two frameworks in reading and language use. The findings contribute to the assessment research literature in that they provide relevant stakeholders with a means of comparing performances on either the LTE (linked to the CEFR scale) or the Chinese CET (linked to the CSE framework).These findings are particularly valuable for western institutions of higher education who, when considering the admission of Chinese students, postgraduate or other, are presented with CET results based on the CSE framework and LTE results based on the CEFR framework.

Keywords: test validation, reading and language use, CEFR, CSE, self-assessment, expert judgement, Rasch measurement

1 Introduction: Exploring comparability between the CEFR and the CSE

The overarching purpose of the research reported in the current paper was to explore how comparability of reading and language use tests such as the LanguageCert Test of English (LTE) and the Chinese College English Test (CET) could be established between the CEFR (the *Common European Framework of Reference for Languages*) and the CSE (the *China Standards of English*) assessment frameworks by using a combination of data from two tests.

The research – which comprised two separate studies – was undertaken because stakeholders needed to be able to compare results using the two assessment frameworks as growing numbers of Chinese students were beginning to be candidates for the LTE. Students in China who have taken the CET, based on the CSE, are nowadays taking the LTE, based on the CEFR, in much greater numbers. Institutions in Europe, including the UK which recognises the LTE for visa and migration purposes, will find the comparisons made in this article valuable in establishing the language proficiency of potential applicants from China, often seeking postgraduate courses, when they apply for admission to courses.

The first study used expert judgement to determine the 'fit' between the CET-type items and the CSE framework, with which the raters were familiar, and the LTE items and the CEFR framework, with which they were less familiar. While there was already existing evidence of the LTE items linkage with the CEFR, the study was expected to establish a reliable linkage between the CET-type items and the CSE framework. The second study was an exercise in which test-takers reported their self-assessments on each of the two tests. This study enabled for a comparison of the accuracy of the test-takers' self-assessments by matching them against their actual test results.

Data for the studies comprised a large cohort (N=2,500) of Year 1 students at a prestigious university in China. These students took two tests and also completed a set of Can Do self-assessments derived from both the CEFR and the CSE. In the first study (Zhao and Coniam 2022), expert judges were asked to place all the CET (College English Test) items on the CSE scale, with which they were very familiar. In the second study (Coniam et al. 2022), students self-assessed their English language ability against CEFR and CSE levels using Can Do statements. The intention of this study was to explore which framework levels students were best able to judge themselves against.

The two studies were originally published separately because there was too much to include in a single paper. The current paper pulls them both together, extending their individual reach, with an attempt to illustrate the issue of comparability from a larger perspective. Descriptions of both studies are consequently essential to provide readers with an adequate understanding of how the findings of the two studies were built upon in a further process of bringing together the analysis of a variety of data to provide a useful and valid picture of how the two scales compare.

2 Background to the CEFR and the CSE frameworks

For the past two decades, the CEFR has come to be accepted as illustrating standard descriptors of language ability by many stakeholders: e.g., policy makers, exam bodies and test developers (Deygers et al. 2018). Not only in Europe, but in many countries around the world (Little 2007), the CEFR has become the common currency for specifying levels of language ability (Figueras 2012).

The CSE reflects an overarching notion of language ability, with which language knowledge and strategies co-function in performing language activities. The CSE development attempts to pull together —in the context of China—a wide range of different English language curriculums and assessment instruments into one overarching framework. The development of the CSE began with the "Common Chinese Framework of Reference for English (CCFR-E): Teaching, Learning, Assessment", which began in 2014 (Jin et al. 2017). This development then became known as the *China Standards of English* (CSE) which was finally released in 2018 and consists of three major levels, each subdivided into three sublevels as illustrated in Figure 1.

Common European Framework of Reference		China Standards of English	
Stage	Level	Level	Stage
Proficient User	C2	Level 9	Advanced
	C1	Level 8	
Independent User	B2	Level 7	
	B1	Level 6	Intermediate
Basic User	A2	Level 5	
	A1	Level 4	
		Level 3	Elementary
		Level 2	
		Level 1	

Figure 1. CEFR and CSE Frameworks

As mentioned above, the test aligned to the CEFR in the current study is derived from the LanguageCert Test of English (LTE). The LTE is a 'level agnostic' test (one that is independent of levels and modes), existing in two parallel formats, both drawn from the same item bank: a paper-based version and an online adaptive test. The validation of the paper-based version of the LTE was reported in . (2021a), and the validation of the adaptive test in Coniam et al. (2021b).

The overarching LanguageCert Item Difficulty (LID) scale is aligned to the CEFR (Coniam 2021a) as illustrated in Table 1. The LID scale has been developed by LanguageCert. It is a validated measurement scale, a necessary prerequisite for any examination/assessment system.

Table 1. ***LID scale***

CEFR level	LID scale range	Mid point
C2	151-170	160
C1	131-150	140
B2	111-130	120
B1	91-110	100
A2	71-90	80
A1	51-70	60

The next section reviews details of previous comparability studies.

3 Relevant comparability studies

One of the first large scale comparability studies was the Cambridge-TOEFL comparability study conducted in the late 1980s (Bachman et al. 1995). This study, which investigated the comparability of the First Certificate of English (FCE) and the Test of English as a Foreign Language (TOEFL), is notable for two reasons. Firstly, it established a baseline for comparability studies and secondly, it initiated comparability studies for different high-stakes tests.

Bachman et al. (1995) set the standard in terms of test taker samples, test types and scoring procedures selected for analysis. Against these robust background measures, conclusions could be drawn about comparability between the two tests: "score comparisons across tests are justified and could be made in a meaningful way" (Bachman 1990: 48).

There have since been numerous studies investigating and establishing comparability between tests. Some of these have been robust studies; some less so, merely claiming comparability. Some of these studies are discussed below.

3.1 Evidence-based equivalence-establishing studies

Detailed recommendations regarding procedures and methods for comparing tests are outlined in the Council of Europe manual (2009) relating language examinations to the CEFR document. According to the manual, establishing large-scale test comparisons requires a considerable amount of data, resources, and analysis, as discussed in the studies below.

Apart from the Bachman et al. study (1995), there have been other large-scale studies that have investigated the comparability of two different tests. Such studies broadly follow procedures described in Bachman et al. (1995); i.e., conducting an analysis of both test content and test results.

Taiwan's General English Proficiency Test (GEPT) assesses learner proficiency across the range of Taiwan's English education framework. A number of robust studies have been conducted, comparing the GEPT with the CEFR, with these studies, for the most part, adhering to the guidelines in the Council of Europe manual (2009), and using expert judgement panels. A brief commentary is provided below.

Brunfaut & Harding (2014) conducted a study investigating the comparability of the GEPT and CEFR listening tests. They concluded that GEPT listening test levels 1-4 largely corresponded to CEFR levels A2 to C1. Green & Inoue (2017) investigated the comparability of GEPT and CEFR speaking test levels. Their analysis indicated that the GEPT speaking tests were generally aligned well with CEFR levels. Knoch & Frost (2016) explored the alignment of GEPT writing tests to the CEFR. While results indicated that the GEPT writing tests aligned with CEFR levels, a slight lowering of GEPT pass scores was recommended in order to better align with CEFR levels.

In a data-driven comparability study, Kunnan & Carr (2017) explored the comparability of GEPT and Internet-Based Test of English as a Foreign Language reading and writing tasks via tests administered to test takers in Taiwan and the USA. They concluded that the two tests were broadly comparable. While the two tests generally assessed the same reading constructs, the reading focus was slightly different in each test.

3.2 Evidence-based CSE/CEFR comparability studies

Alderson (2017) discussed a range of studies exploring the CSE and its correspondence to the CEFR. These have been augmented by the work of Jin et al. (2017) and Zhao et al. (2017), investigating the linking of College English vocabulary levels with the CEFR.

Further studies were conducted, showing comparability between the CSE and CEFR. Dunlea et al. (2019) describe a comprehensive study involving all four language skills that explored the relationship between the British Council's Aptis test and IELTS with the CSE. The methodology involved expert judgement of items against CSE and CEFR levels and the assignment of CSE descriptors against tasks. Following this, the proposed levels were field tested in an "external evaluation" exercise, where Chinese teachers rated their own students against the proposed matched levels, as illustrated in Table 2 below.

Table 2. ***Level match between the CSE and the CEFR (Dunlea et al. 2019)***

CSE	CEFR
L9	C2
L8	C1
L6/7	B2

L4/5	B1
L3	A2
L2	A1

Peng et al. (2021) report on a study attempting to establish level correspondences between CEFR and CSE levels using difficulty estimates of all published descriptors (467 for the CEFR and 1,051 for the CSE) of ratings by English language teachers and students. While there was close correspondence at the top and bottom ends of the scale, there was overlap in the middle levels. Table 3 elaborates.

Table 3. ***Level match between the CSE and the CEFR (Peng et al. 2021)***

CSE	CEFR
L9	C2
L7/L8	C1
L6/L7	B2
L4/L5	B1
L2/L3	A2
L2	A1
L1	A0

As may be seen, while there is a good degree of agreement in the correspondence between the two studies that tested all four skills, there are also divergences. These may be due to a number of factors: the samples, the tests, the judges used in the ratings. These factors will, over time, be investigated by studies that focus on one factor at a time so that, in spite of a good degree of agreement, the divergences may also be explored.

3.3 Comparability studies with little supporting evidence

The discussion above has reported on studies claiming evidence of comparability between tests. Comparability has been claimed for many tests, usually with the CEFR. Often, such claims have been made – and are still made – on the basis of little or no apparent evidence – see Table 4 below.

The (now inaccessible) AWEMAP project from the early 2000's laid out numerous tables indicating apparent equivalence[1]. As Green (2012) comments, however, on certain of AWEMAP's equivalences, "convincing evidence of the relationship" for certain scales was simply not available (2012: 87). Such 'mappings' listed by AWEMAP's included Ordinate's Phone Pass, mapped to the CEFR Scale by Ordinate; DynEd Dynamic Education's Placement Test, mapped to the ILR OPI Scale and TOEIC by DynEd. The latter, mapping "exam correlations", can be found at https://www.dyned.com/media-library/correlations-intl/.

Current claims about comparability, usually with the CEFR scales – for which no evidence is provided —may be seen in the claims of numerous education bodies. Table 4 presents a small sample of specious claims.

1. AWEMAP was the Worldwide English LET EFL ESL EALLEP ESOL Assessment Scales and Tests Mapping Project. It was last available at http://www.geocities.com/esolscale/index.html?200510.

Table 4. ***Apparent mappings to the CEFR***

Education Body	URL
Express Publishing's Vocational English Certificate (VEC)	http://ecahe.eu/w/index.php/English_language_test_equivalency_table
EF Education First's Standard English Test (EF SET)	https://www.ef.com/wwen/english-tests/test-comparison/score-converter/

4 Overarching purpose: Exploring comparability between the CEFR and the CSE

As stated in the abstract, the overarching purpose of the research conducted in the two studies reported here involved exploring how comparability in reading and language use tests (such as the LanguageCert Test of English) might be established between the CEFR and the CSE.

4.1 Data

A variety of test taker/self-assessment/test/expert judgement data was collected from a Chinese university from late 2020 to May/June 2021. The principal focus was Year 1 CET students. Given that the university admits a considerable number of overseas Chinese students, it could be taken that there would be a considerable student ability spread – ranging across the CEFR and CSE. Table 5 presents a picture of the data collected.

Table 5. ***Project data***

Sample	Instrument	Timeframe
2,498 Year 1 CET students	65-item in-house CET Reading and Language Use placement test	Oct 2020
4,128 Year 1 CET students	53-item LTE Reading and Language Use test	Mar 2021
4,128 Year 1 CET students	16 CEFR self-assessment Can Do statements	May 2021
4,128 Year 1 CET students	22 CSE self-assessment Can Do statements	May 2021
8 ESL university professors	Expert judgement of 23 discrete-point LTE items	May 2021
8 ESL university professors	Expert judgement of 30 discrete-point CET items	Jun 2021
2,311 Year 1 students	Official CET 4 test	Oct 2021

In late 2020, approximately 2,500 Year 1 CET students took a 65-item multiple-choice reading and language use test prepared by experts from the university. Approximately three months later, this same set of students took a 53-item multiple-choice LanguageCert reading and language use test constructed from the LTE item bank. The LTE items were selected on the basis of having been calibrated to represent the spectrum of difficulty across the six CEFR levels. The items had been adapted from the material validated in Coniam et al. (2021a). The composition of the tests is described in the following section.

4.2 Content analysis of the tests

This section presents a comparative analysis of the make-up of the two reading and language use tests. Table 6 elaborates.

Table 6. *Component Analysis of CET and LTE Tests*

CET	LTE
Cloze: 15 items One 15-item cloze passage assessing grammar, syntax, discourse, vocabulary	Cloze: 30 items Three 5-item cloze passages assessing grammar, syntax, discourse, vocabulary
Discrete items: 30 items 30 items assessing grammar, syntax, vocabulary, language use	*Discrete items: 23* items 23 items assessing grammar, syntax, vocabulary, language use
Reading comprehension: 20 items Four 5-item reading comprehension passages assessing a range of reading comprehension skills	Reading comprehension: 15 items Three 5-item reading comprehension passages assessing a range of reading comprehension skills
65 items	53 items

As Table 6 illustrates, the CET test is slightly longer than the LTE test; also, all CET items were four-option multiple-choice whereas the LTE items were three-option multiple-choice. Despite these minor differences, the content of the two tests, and even the order in which the different sections of the test were presented to test takers, exhibit a broad amount of similarity.

4.3 Test administration and expert judge study

Test takers took the CET 65-item test in late 2020 as part of their university course whereas the official CET is usually taken at the end of the academic year.

In 2021, the same group of students took the LanguageCert 53-item test. Test takers subsequently completed two Can Do self-assessment profiles. One profile consisted of 16 Can Do statements drawn from the CEFR and the second 22 statements drawn from the CSE. The composite set of 38 items were all presented bilingually in both English and Chinese, with CEFR and CSE items and levels intermingled in an attempt to reduce the chances of respondents trying to guess where their own estimated ability level finished.

The focus for the expert judge study was the discrete items in the two tests. There were 30 such items in the CET test and 23 items in the LTE test, testing grammar, syntax, vocabulary and language use: the second component of the test presented in Table 6 above. There were eight expert judges, professors from the Foreign Studies Department, all of whom had been involved in setting CET items for their students at the university.

Before rating took place, training and standardisation sessions were conducted for the expert raters participating in the study. The purpose of these sessions was to increase rater reliability and familiarity with the less known CEFR framework. First, they rated sample CET items using the nine-level CSE. They then rated sample CEFR items using the six-level CEFR. Following this, the expert raters were given the 30 CET items to rate against the nine CSE levels and the 23 LTE items against the six standard CEFR Levels.

5 Statistical analysis

In the current study – to gauge test fitness for purpose, and to link two different tests to a common scale —both Classical Test Statistics (CTS) and Rasch measurement have been used and are briefly outlined below. CTS analysis reports test mean, standard deviation and test reliability. Rasch measurement facilitates the calibration of different facets within and between tests.

5.1 Classical Test Statistics (CTS)

The test mean for a proficiency test tends to be within a range of 60-70% (Heaton 1990)[2]. This will depend, however, on where the pass mark is set by the exam body concerned, and the purpose for which the test is intended. A test mean of around 60-70% suggests that the test is generally appropriate to the level of a 'typical' test taker (Burton et al. 1991). Such a mean in general indicates that most test takers managed to finish the test and that test takers may be assumed to have done their best.

In terms of test reliability – where levels of reliability are associated with test length (Ebel 1965) – a desirable level is generally taken as 0.7 for tests with 65 or more items.

5.2 Rasch measurement

The use of the Rasch model enables different facets to be modelled together, converting raw data into measures which have a constant interval meaning (Wright 1997). This is often likened to measuring length using a ruler, with the units of measurement in Rasch analysis (referred to as 'logits') evenly spaced along the ruler. In Rasch measurement, test takers' theoretical probability of success in answering items is gauged – scores are not derived solely from raw scores. While such 'theoretical probabilities' are derived from the sample assessed, they are able to be interpreted independently from the sample due to the statistical modelling techniques used. Measurement results based on Rasch analysis may therefore be interpreted in a general way (like a ruler) for other test taker samples assessed using the same test. Once a common metric is established for measuring different phenomena (test takers and test items in the current instance), test taker ability may be estimated independently of the items used, with item difficulties also estimated independently from the sample (Bond et al. 2020).

Since test taker measures and item difficulties are placed on an ordered continuum in Rasch, direct comparisons between test taker abilities and item difficulties, as mentioned, may then be conducted, with results able to be interpreted with a more general meaning. One of these more general meanings involves the transferring of values from one test to another via anchor items. Anchor items are a number of items that are common to both tests; they are invaluable aids for comparing students taking different tests; and were used in the current study. Once a test, or scale, has been calibrated, the established values can be used to equate different test forms.

In the current study, the LTE test has been compiled, as mentioned, from robust test material validated in Coniam et al. (2021a).

5.3 CTS analyses

CTS item analyses are presented in Table 7. Results for the whole test are presented first for all test takers. Second, since expert judgement of difficulty was judged against the sets of discrete items from both tests, an analysis of the exact-same set of test takers is also presented for purposes of direct comparison.

A 'good' item is defined by Falvey et al. (1994) as one with a facility index of 30%-80%. A 'reliable' item is defined by Falvey et al. (1994) is one with a discrimination index greater than 0.3.

2. Heaton (1990) states, in the case of means, that with proficiency tests, test scores should ideally be spread out "over the whole range of the scale" (p. 171) (i.e., with a mid point of around 0.6), which aids in discriminating among test takers. The corollary is that proficiency tests will have a slightly lower mean than classroom tests. This is reflected in, for example, the mean performance score on key proficiency tests. To exemplify, for 2019, the overall mean on IELTS Academic was 6.08/9 and IELTS General Training 6.59/9. Cf., https://www.ielts.org/for-researchers/test-statistics/test-taker-performance.

Table 7. *Item analyses*

	CET Whole test	LTE Whole test
Test takers	2,498	4,128
Items	65	53
Mean	57.31 (60.3%)	28.31 (53.4%)
SD	12.1 (12.7%)	6.6 (12.5%)
Reliability KR20	0.83	0.72
Good items	39/53 (74%)	51/65 (78%)
	CET Discrete items	LTE Discrete items
Test takers	2,318	2,492
Items	30	23
Mean	15.71 (52.3%)	13.15 (57.2%)
SD	3.9 (13.1%)	3.38 (14.7%)
Reliability KR20	0.63	0.62
Good items	21/30 (70%)	20/23 (87%)

As Table 7 indicates, whole test analyses were broadly comparable. Standard deviations and reliability (as measured by the Kuder-Richardson KR20 statistic) were very close on both tests. Both test means were in the 'desirable' range – in the 50-60% range, suggesting that the tests broadly fit the target population, and that most test takers finished the test and had given it their best shot. Test reliability for both tests was close to or above 0.8, indicating that the tests may be assumed to have been well constructed.

The more focused picture with the discrete items showed an even closer match between the two tests, suggesting that comparisons may be seen as generally valid.

While CTS gives a baseline indication of comparability, if tests are to be linked so that they may be directly compared, Rasch measurement needs to be applied, because only then can analysis be carried out to determine comparability between two non-linked and separate tests.

5.4 Data and frame of reference

To recap, there are four sets of assessment data in the current study: the 65-item CET test, the 53-item LTE test, 22 CSE-referenced Can Do ratings and 16 CEFR-referenced Can Do ratings. Since all four datasets were collected from the same test takers, the data configuration may be taken as a unified collection, in that all data are referenced to the same candidates and to their English language ability. The *person links* (Boone 2016) in the four datasets embrace a coherent *frame of reference* (FOR) (Humphry 2006)[3].

In order to calibrate the four datasets in the current study onto the LanguageCert Item Difficulty (LID) scale (Table 1), a previously calibrated test (henceforth referred to as the "anchor test") from the Coniam et al. (2021a) study was incorporated into the data. As a subset of the anchor test, the LTE test in the current study provides a set of *item links* (Boone 2016). With sets of both person links and item links established, the LTE test could then be linked to the anchor test. Following this, the other datasets in the study – the CET test and the two sets of self-assessments – could then be calibrated against the anchor

3. Humphry (2006) defines a frame of reference as "compris[ing] a class of persons responding to a class of items in a well-defined assessment context." In a given frame of reference, a number of disparate datasets may then be calibrated together and aligned to each other.

test onto the LID scale. This resulted in all five assessment datasets being included into one single FOR.

For analysis and calibration purposes, 100 has been taken as the mid-point of the scale (see Table 1 above). To this end, Rasch logit values are rescaled to a mean of 100 and a standard deviation of 20.

5.5 Single frame of reference analysis

As mentioned, the *anchor test* had been previously anchored to the LID scale. Against this backdrop, the composite analysis is presented in Appendix 1.

To recap, item links in the overall dataset were established between the 53 items in the LTE test and the anchor test. Person links were established via the two tests and the two sets of self-assessments. All five datasets may therefore be seen to be within an overall frame of reference – the composite analysis to the far left of the person-item map in Appendix 1. In the analysis of the data, the two tests fit the Rasch model well, with mean square infit and outfit figures within the 0.5 to 1.5 range, and high reliability. The means for both tests were very comparable, both approximately a quarter of a logit above the overall mean of 100. The two tests emerged as being of comparable difficulty, if slightly more demanding than the mean calibration point. In contrast, respondents tended to slightly overestimate their abilities on both the CEFR and the CSE, with self-assessment mean values slightly higher than their actual results indicated.

5.6 External test reference point

Official CET-4 scores were obtained in November 2021 for 2,311 of the test takers. Table 8 presents the Pearson correlations between the official CET-4 test for reading, the LTE reading and language use (RLU) test and the China university CET reading and language use test used in the current study.

Table 8. ***Pearson correlations in CET tests***

Test	Correlation detail	Official CET-4 reading test	CET RLU test
LTE RLU test	r	0.71	0.73
	p level	< .001	< .001
CET RLU test	r	0.74	
	p level	< .001	

As can be seen, the tests inter-correlate significantly at the 0.7 level. The highest correlation, as might be expected, is between the official CET-4 reading test and the in-house CET RLU test at 0.74. Given that the accepted correlation for tests of 65 items is around 0.7 (Ebel 1965), the inter-test correlations in the current study are an indication that the tests are broadly assessing similar constructs.

In sum, it may therefore be seen that the two tests are similar in construction. Test means, standard deviations and reliability figures were broadly comparable at two levels: at the whole test level and on the reading and language use subtests. The two cloze subtests correlate at the 0.7 level, with the shared variance overlap indicating that the two cloze subtests are potentially measuring approximately 50% of the same construct.

6 The two studies

As mentioned, two studies were undertaken in addition to the large-scale testing of candidates on a CET-type and LTE-type test. These will now be further discussed as a basis for subsequently triangulating the results of those two formal tests. Without a process of triangulation, it is difficult to accurately determine comparisons between the two separate tests.

The first study involved a set of China English language professors in expert rating. The professors first rated the CET reading and language use items against the CSE scales; second, they rated the LTE reading and language use items against the CEFR levels.

The second study involved a series of self-assessment Can Do statements describing English language competences. Test takers who took the two tests later self-evaluated using the Can Do statements after finishing the two tests. One set of Can Do statements generated self-assessments based on the CSE levels, while the other set generated self-assessments based on the CEFR levels. The use of instruments such as Can Do statements in self-assessment has been validated in a number of studies (see e.g., Brown et al. 2014; Summers et al. 2019). These two datasets were analysed along with the two tests.

In order to assist the reader, key issues in and outcomes from the two studies are reproduced in the following section.

7 Study one: Expert judge ratings

The overarching hypothesis in the study was that levels of agreement achieved by expert judges rating the CET items against the CSE – with which they were very familiar – would be better than levels of agreement achieved rating LTE items against the CEFR – with which they were less familiar. The research question pursued in the study was:

(RQ1) To what extent are expert judges more accurate in their judgement of item difficulty when rating test items against a framework with which they are very familiar, as opposed to rating test items against a framework with which they are less familiar.

Against this backdrop, a high level of agreement between student test scores and expert-rated values was hypothesised for the CET items (i.e., a 'strong agreement' [0.8] in Kappa statistic terms (Landis and Koch 1977). Conversely, with the LTE items, only a moderate level of agreement ('substantial agreement' [0.6] in Kappa terms) was hypothesised between student test scores and expert-rated values.

The study involved eight expert judges, professors from the Foreign Studies Department, all of whom had set CET items for their students at the university. Given the relevance and status of the CSE in China, the eight expert judges had a clear picture of standards in the CSE, confirmed by senior staff at the university. Given the fact that they were all English language professionals, most had also knowledge of, albeit not in-depth familiarity with, the CEFR.

In standardisation sessions, the eight judges trial-rated sample CET items against the nine CSE levels, and sample CEFR items against the six CEFR levels. The judges then rated the 30 discrete CET items against the CSE and the 23 discrete LTE items against the CEFR.

Following the training and standardisation, test taker mean scores and expert judge mean ratings of the discrete items for each cloze subtest were equated by aligning the differences between the means and standard deviations of both sets of scores. Table 9 presents the findings which emerged following analysis of test takers' scores on the tests and from judges' ratings of item difficulty. Both tests, as mentioned, were anchored at 100 – the mid-point of the LanguageCert scale at which all LanguageCert tests are anchored (see Lee et al. 2022).

Table 9. ***Test taker mean scores and expert judge mean ratings of CET and LTE items***

Subtest type and mean	Items	Mean LID value	SD	Reliability
LTE test taker mean	23	102.96	32.10	0.97
LTE expert mean rating	23	104.89	32.77	1.00
CET test taker mean	30	104.28	22.43	1.00
CET expert mean rating	30	96.25	20.22	0.84

As a baseline, verification of reliability and Rasch fit statistics were first conducted. Reliabilities for all four elements of the dataset were high, above 0.8. Rasch infit and outfit figures were within acceptable levels (0.5-1.5).

With the LTE 23 discrete LTE items, the test taker mean score was 102.96; the expert judge mean rating was 104.89, a difference of 1.93.

With the 30 discrete CET items, the test taker mean score was 104.28; the expert judge mean rating was 96.25, a difference of 8.03. Differences between test taker mean scores and expert judge mean ratings were less than 10 points, the half-logit difference generally accepted as non-significant (Zwick et al. 1999).

Standard deviations (SD) were broadly comparable within each pair of tests. However, as was the case with the mean scores, the SDs differed between tests. To smooth out these differences, means and SDs needed to be aligned into a single frame of reference (see Linacre 2022). Following the 'smoothing-out', or alignment, directly comparable values were then able to be computed for each LTE and CET item by subtracting the expert judge rated item mean from the test taker item mean.

With all LTE and CET items in the same frame of reference, expert judge-rated and test taker mean item values could finally be mapped to CEFR levels. On both tests, for each item, the CEFR/CSE level match between the expert judge mapped level and the test taker score level was examined. A tally was made of whether the match was exact, or whether the difference was lenient or strict by one or more CEFR levels. The results are presented in Table 10.

Table 10. ***Fit of expert judge mapped levels to test taker scores levels***

	CET items	LTE items
Number of items	N=30	N=23
Over-rated by one level	0	0
Exact fit	27 (90.0%)	5 (21.7%)
Under-rated by one level	3 (10.0%)	18 (78.3%)
Kappa	0.92 (p<.001)	0.40 (p<.001)

With the CET items, 27/30 (90%) of the expert ratings matched the test taker mean score values; three items were under-rated by one level. With the LTE items, however, expert ratings matched test takers' scores much less closely than was the case with the CET items. Only 5/23 (21.7%) of the expert ratings on the LTE items matched test takers' scores on the LTE items. 18 items were under-rated by one level. This indicates that the expert judges do not have as good an understanding of the CEFR as they do the CSE levels, hence the under-rating of many items against the CEFR levels. The implications that can be drawn from these findings are that, in future studies and practices of comparability, additional standardisation and training should be given to raters in the frameworks with which they are less familiar.

After recording expert judge-mapped levels against test taker mean score levels (as 1-6, where A1=1 and C2=6), Kappa was calculated, with the results presented in the final row of Table 10. With the CET items, a Kappa of 0.92 (p<.001) emerged – a 'strong' agreement between the two variables. With the LTE items, a Kappa of 0.40 (p<.001) emerged between the two variables – only a 'fair' agreement.

The conclusion drawn from the expert rating study was that judges who are very familiar with their own assessment situation in terms of test material, test constructs, assessment levels etc. are able to make more accurate assessments than are judges who are less familiar with the material they are assessing, and the levels at which test items should be assessed. While the results in the expert rating study might appear to be somewhat self-evident, the results lend support to the argument that expert judgement is a methodology that may be reliably utilised in test validation provided the raters are completely familiar

with and experts in that specific test. This lends credence to the fact that the assessments in the current study may be taken as reliable. One implication is that raters who are less familiar with one of the tests require further standardisation and training.

8 Study two: Test taker self-assessment study

Against the backdrop outlined above, the second published study, an essential component of the triangulation that eventually took place, pursued two questions.

(RQ1) To what extent can self-assessments be validly used to establish correspondences between the CEFR and CSE frameworks?

(RQ2) To what extent are correspondences between the CEFR and CSE frameworks in line with those reported in previous studies?

Subsequent to having taken both tests, test takers responded to 38 Yes/No-framed Can Do statements. 22 Can Do statements related to the CSE and 16 Can Do statements to the CEFR. The results were analysed in the same frame of reference as the two tests, where the anchor point was 100, the LID scale mean. Table 11 presents the results for the test means and the self-assessment means. Appendix 1 presents the big picture.

Table 11. ***Test taker test self-assessment mean scores***

Subtest type and mean	Items	Mean LID value
LTE test mean	53	105.33
CET test mean	65	104.10
CEFR Can Do assessment mean	16	95.29
CSE Can Do assessment mean	22	95.66

As can be seen, the means of both two sets of self-assessments are comparable. The fact that both sets of self-assessments are five LID scale points (a quarter of a logit) below the anchored mean of 100 is indicative of test takers tending to slightly over-rate themselves – a not uncommon phenomenon (Dunning 2006).

The fact that means for both tests and self-assessment ratings were acceptably within half a logit (Zwick et al. 1999), suggested that test takers could be considered sufficiently objective in their self-assessments to permit tentative correspondences to be drawn between CSE and CEFR levels. Correspondences were then drawn up between the two sets of self-assessments for each CEFR level. To exemplify, Table 12 below presents the CEFR and CSE Can Do statements for the LanguageCert B2 level (111-130 LID scale points).

Table 12. ***CEFR and CSE Can Do statement level comparison chart: B2 (111-130)***

CEFR			CSE		
CEFR Can Do Statements	*LID value*	*CEFR level*	*CSE level*	*LID value*	CSE Can Do Statements
			L7	129.72	I can understand linguistically complex novels and materials related to culture and appraise their linguistic features.

CEFR			CSE		
			L6	128.73	I can understand the terminology of operational texts in related professional areas.
			L7	127.85	I can understand book reviews in relevant fields of inquiry.
			L6	127.27	I can understand novels and argumentative texts comprised of relatively complex language.
I can scan through rather complex texts, e.g. articles and reports, and can identify key passages.	118.74	B2			
			L5	117.63	I can understand the common figures of speech in stories pertaining to social life written in relatively complex language.
I can understand in detail specifications, instruction manuals, or reports written for my own field of work	116.58	B2			
			L5	116.41	I can infer the content of an entire book or text by scanning the table of contents.
I can read texts dealing with topics of general interest, such as current affairs, without a dictionary, and can understand multiple points of view.	115.69	B2			

Within the B2 CEFR LID range of 111-130, three CEFR C1 self-assessments were found, along with six CSE self-assessments, of which two were at L5, two at L6 and two at L7.

The B2 CEFR / CSE fit was therefore interpreted as being CEFR level B2 fitting quite broadly against CSE levels L5-L7.

Conclusions regarding RQ1 were that, while respondents tended to slightly overestimate their abilities on both the CEFR and the CSE, such overestimations were minimal, in that mean values were only a quarter of a logit higher than might have been expected. Overestimations were also consistent with the scales for both frameworks. The premise that self-assessments could be used to establish correspondences between the CEFR and CSE frameworks was then accepted.

Regarding RQ2, correspondences which emerged between the CEFR and CSE frameworks were broadly in accordance with those proposed by previous studies. While there were some divergences,

more notably towards the lower end of the scales, the correspondences proposed broadly echo those reported in previous studies.

The results reveal that – with the exception of CEFR level C2, for which there was insufficient data to perform a calibration, it was possible to produce an overall tentative mapping of how the CEFR scale, as represented by the LTE, might be mapped against the CSE scale as represented by the CET-based assessment. Table 13 presents the fit.

Table 13. ***Current study CEFR / CSE fit***

CEFR	CSE
C2	N/A
C1	L7
B2	L5-L7
B1	L4-L6
A2	L3-L5
A1	L2-L3
A0	L1

As can be seen from Table 13, while there is not a one-to-one match between the levels in the two frameworks, as one moves up the scale, there is a graduated fit between the CEFR and the CSE.

Table 14 below extends the picture of alignments presented in Table 13, and includes the alignments proposed in Dunlea et al. (2019) [the 'Dunlea' study] and in Peng et al. (2021) [the 'Peng' study].

Table 14. ***Extended CEFR / CSE mapping***

Current study		*Dunlea study*		*Peng study*	
Reading & Language Use		All skills		All skills	
CSE	*CEFR*	CSE	CEFR	*CSE*	*CEFR*
	C2	L9	C2	L9	C2
L7	C1	L8	C1	L7-L8	C1
L5-L7	B2	L6-L7	B2	L6-L7	B2
L4-L6	B1	L4-L5	B1	L4-L5	B1
L3-L5	A2	L3	A2	L2-L3	A2
L2-L3	A1	L2	A1	L2	A1
L1	A0	L1		L1	A0

The different mappings revealed both similarities and differences. For readability sake, these are listed below.

- The current study mapped A0 onto L1, as did the Peng study.
- The current study mapped A1 against L2/L3. The Dunlea study mapped A1 to L2, and the Peng study mapped A1 to L2.
- The current study mapped A2 more broadly against L3-L5. The Dunlea study mapped A2 to L3 while the Peng study mapped A2 against L2/L3.
- The current study mapped B1 against L4/L6. The Dunlea and Peng studies mapped B1 against L4/L5.

- The current study mapped B2 against the bottom end of L5 to L7. The Dunlea study mapped B2 against L6/L7 and the Peng study mapped B2 against L6/L7.
- The current study mapped C1 at L7. The Dunlea study mapped C1 at L8 while the Peng study mapped C1 at L7/L8.
- There was no data for C2 in the current study.
- The results of the current study can therefore be seen to broadly reflect the mappings of the previous Dunlea et al. (2019) and Peng et al. (2021) studies. As mentioned above and shown in Table 14, both Dunlea et al (2019) and Peng et al 2021) recently published the results of comparative studies. These studies mapped the results of all four skills at appropriate levels within the CEFR and CSE frameworks. They are useful because they reveal both similarities and differences in the findings. The contribution of the current study is that it too has mapped results against the CSE and CEFR frameworks. These findings broadly reflect the mappings of both Dunlea et (2019) and Peng et al (2021) so we have a broader perspective of these comparison studies.

9 Conclusion

The paper has investigated comparability between the CEFR and the CSE. Comparisons have been explored from the perspective of tests of reading and language use, as produced by LanguageCert for the CEFR and a comparable test of reading and language use from a China CET (College English Test) produced by a university in China.

Datasets in the study comprised a large cohort (N=2,500) of Year 1 students at a prestigious university in China. These students took two tests of reading and language use: one, initiated, developed, trialled and standardised by ESL professors within the university and the other derived from the LanguageCert Test of English. Analyses of both whole tests and reading and language use sections were seen to be broadly comparable. Classical Test Statistics indicated that both tests were reliable, with comparable means and standard deviations. A Rasch analysis showed that the two tests fit the Rasch model well, with acceptable mean square infit and outfit figures, and high reliability figures. Inter-test correlations between the two tests and the official CET-4 test results emerged in the 0.7 range. Such a level of correlation is considered reasonable, indicating that it is possible to come to conclusions about the amount of comparability between the two different tests.

Against this backdrop of two comparable and reliable tests, two different studies were undertaken, to investigate CEFR/CSE comparability issues.

In the first study, expert judges rated CET items against the CSE, with which they were very familiar, and LTE items against the CEFR, with which they were not as familiar. A high level of agreement between test taker mean score values and expert-judge-rated values on the CET discrete items was obtained. This suggested that values obtained between the cross-calibration of the CET and LTE tests could be seen as robust, and could support the other analyses conducted.

In the second study, test takers completed sets of Can Do self-assessments related to CEFR and CSE scale levels. Test takers tended to slightly overestimate their abilities, and although there was not a one-to-one match between the levels in the two frameworks, correspondences between the CEFR and CSE frameworks were nonetheless broadly in accordance with those proposed by previous studies (Dunlea et al. 2019; Peng et al. 2021).

The implications for stakeholders (students, teachers, administrators and universities) are that broad comparisons may be seen between CSE-aligned tests and CEFR-aligned tests. The mapping for reading and language use between the CEFR and the CSE as it emerged in the current study was reported in Table 13 above.

As reported earlier, as one moves up the scale, there is a graduated fit between the CEFR and the CSE. While there are some divergences, more notably towards the lower end of the scale, the correspondences

broadly echo those reported in previous studies. This study therefore provides a useful first stage in the comparison between the LanguageCert Item Difficulty scale and, hence, LanguageCert tests, and the CET-type tests linked to the China Standards of English. It also contributes to the research literature on how comparability between separate non-linked tests can be investigated and established. In summary, both a theoretical and practical contribution have been made by this study: theoretically, the findings add to the knowledge provided by different comparison studies of these two important assessment frameworks and the exams that have been centred within them; practically, students in China who take both the CET and the LTE can use those results to inform and enlighten admission tutors in both the UK and the rest of Europe.

10 References

Alderson, J. Charles. 2017. Foreword to the special issue "the common European framework of reference for languages (CEFR) for English language assessment in China" of language testing in Asia. Language Testing in Asia, 7(1), 1-9.

Bachman, Lyle. F. 1990. A comparison of the abilities measured by the Cambridge and Educational Testing Service EFL test batteries. *Issues in Applied Linguistics 1*(1). 30-54.

Bachman, Lyle F., Fred Davidson, Katherine Ryan & Inn-Chull Choi. 1995. *An investigation of the comparability of the two tests of English as a foreign language: the Cambridge-TOEFL comparability study*. Cambridge: Cambridge University Press.

Bond, Trevor, Zi Yan & Moritz Heene. 2020. *Applying the Rasch model: Fundamental measurement in the human sciences*. Routledge. https://doi.org/10.4324/9780429030499.

Boone, William. J. 2016. Rasch analysis for instrument development: why, when, and how? *CBE-Life Sciences Education* 15(4), rm4. https://doi.org/10.1187/cbe.16-04-0148.

Brown, N. Anthony, Dan P. Dewey & Troy L. Cox. 2014. Assessing the validity of can□do statements in retrospective (then-now) self-assessment. *Foreign Language Annals 47*(2). 261-285. https://doi.org/10.1111/flan.12082.

Brunfaut, Tineke & Luke Harding. 2014. *Linking the GEPT listening test to the Common European Framework of Reference*. Taipei, Taiwan: Language Training and Testing Center.

Burton, Steven J., Richard R. Sudweeks, Paul F. Merrill & Bud Wood. 1991. *How to prepare better multiple-choice test items: Guidelines for university faculty*. Brigham Young University Testing Services and The Department of Instructional Science: Utah. https://testing.byu.edu/handbooks/betteritems.pdf.

Choi, Inn-Chull & Youngsun Moon. 2018. A comparability study of two standardized English as a foreign language tests. (Language Research. https://doi.org/10.30961/lr.2018.54.2.277).

Choi, Inn-Chull. 1995. A comparability study on SNUCREPT and TOEIC. (Language Research).

Coniam, David, Tony Lee, Michael Milanovic & Nigel Pike. 2021a. Validating the LanguageCert Test of English scale: The paper-based tests. London, UK: LanguageCert.

Coniam, David, Tony Lee, Michael Milanovic & Nigel Pike. 2021b. Validating the LanguageCert Test of English scale: The adaptive test. London, UK: LanguageCert.

Coniam, David, Wen Zhao, Tony Lee, Michael Milanovic & Nigel Pike. 2022. The role of expert judgement in language test validation. *Language Education and Assessment, 5*(1), 18–33.

Council of Europe. 2009. *Relating language examinations to the Common European Framework of References for Languages: Learning teaching, assessment*. Strasbourg: Language Policy Division.

Deygers, Bart, Koen Van Gorp & Thomas Demeester. 2018. The B2 level and the dream of a common standard. *Language Assessment Quarterly 15*(1). 44-58. https://doi.org/10.1080/15434303.2017.1421955.

Dunlea, Jamie, Richard Spiby, Sha Wu, Jie Zhang & Mengmen Cheng. 2019. *China's standards of English language ability: Linking UK exams to the CSE*. British Council. https://www.britishcouncil.org/sites/default/files/linking_cse_to_uk_exams_5_0.pdf.

Dunning, David. (2006). Strangers to ourselves. The Psychologist, 19(10), 600-603.

Dyned. 2022. Exam correlations. https://www.dyned.com/media-library/correlations-intl/.

Ebel, Robert L. 1965. *Measuring educational achievement*. Prentice-Hall, NJ: Englewood Cliffs.

Falvey, Peter, Jack Holbrook & David Coniam. 1994. *Assessing students*. Hong Kong: Longman.

Figueras, Neus. 2012. The impact of the CEFR. *ELT Journal 66*(4). 477-485. https://doi.org/10.1093/elt/ccs037.

Green, Anthony & Chihiro Inoue. 2017. *Relating the GEPT Speaking Tests to the CEFR*. Taipei, Taiwan: Language Training and Testing Center.

Green, Anthony. 2012. *Language functions revisited: Theoretical and empirical bases for language construct definition across the ability range (Vol. 2).* Cambridge: Cambridge University Press.

Gu, Mini. 2018. An introduction to China's college English test (CET). *World Education News and Reviews*. https://wenr.wes.org/2018/08/an-introduction-to-chinas-college-english-test-cet.

Heaton, J. B. 1990. Writing English language tests. Harlow, UK: Longman.

Humphry, Stephen. 2006. *The impact of differential discrimination on vertical equating. ARC report*. Western Australia: Department of Education & Training.

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). 1-19. https://doi.org/10.1186/s40468-017-0032-5.

Knoch, Ute & Kellie Frost. 2016. *Linking the GEPT Writing Sub-test to the Common European Framework of Reference.* Taipei, Taiwan: Language Training and Testing Center.

Kunnan, Antony John & Nathan Carr. 2017. A comparability study between the General English Proficiency Test-Advanced and the Internet-Based Test of English as a Foreign Language. *Language Testing in Asia 7*(1). 1-16. https://doi.org/10.1186/s40468-017-0048-x.

Landis, J. Richard & Koch, Gary G. 1977. The measurement of observer agreement for categorical data. Biometrics, 33, 159-174. https://doi.org/10.2307/2529310.

Lee, Tony, Michael Milanovic & Nigel Pike. 2022. Equating Rasch values and expert judgement through externally-referenced anchoring. *International Journal of TESOL Studies* 4(1). 96-112. https://doi.org/10.46451/ijts.2022.01.12.

Linacre, John M. 2012. *A user's guide to WINSTEPS*. Chicago, IL: Winsteps.com.

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. https://doi.org/10.1111/j.1540-4781.2007.00627_2.x.

Peng, Chuan, Jianda Liu & Hongwen Cai. 2021. Aligning China's Standards of English language ability with the Common European Framework of Reference for Languages. Asia-Pacific Education Researcher. Springer. https://doi.org/10.1007/s40299-021-00617-2

Summers, Maria M, Troy L. Cox, Benjamin L. McMurry & Dan P. Dewey. 2019. Investigating the use of the ACTFL Can Do statements in a self-assessment for student placement in an Intensive English Program. *System 80.* 269-287. https://doi.org/10.1016/j.system.2018.12.012.

Wright, Benjamin. D. 1997. A history of social science measurement. *Educational Measurement: Issues and Practice* 16(4). 33-45. https://doi.org/10.1111/j.1745-3992.1997.tb00606.x.

Zhao, Wen, Boran Wang, David Coniam, & BIngxue Xie. (2017). Calibrating the CEFR against the China Standards of English for College English vocabulary education in China. *Language Testing in Asia* 7(1), 1-18. https://doi.org/10.1186/s40468-017-0036-1.

Zhao, Wen & David Coniam. (2022). Using Self-Assessments to Investigate Comparability of the CEFR and CSE: An Exploratory Study Using the LanguageCert Test of English. International Journal of TESOL Studies, 4(1), 169-186. doi.org/10.46451/ijts.2022.01.11.

Zwick, Rebecca, Dorothy T. Thayer & Charles Lewis. 1999. An empirical Bayes approach to Mantel-Haenszel DIF analysis. *Journal of Educational Measurement 36*(1). 1-28. https://doi.org/10.1111/j.1745-3984.1999.tb00543.x.

11 Biographies

David Coniam [corresponding author] is Head of Research at LanguageCert. He has been working and researching in English language teaching, education and assessment for almost 50 years. His main publication and research interests are in language assessment, language teaching methodology and academic writing and publishing.

Michael Milanovic is Chairman of LanguageCert and a member of its Advisory Council. Previously CEO of Cambridge Assessment English, he has been working extensively with PeopleCert since 2015. He worked closely with the Council of Europe on its Common European Framework of Reference, has held, and still holds a number of key external roles.

Wen Zhao is Dean of the School of Foreign Studies at Jinan University, Guangzhou. Her main publications and research interests are in corpus linguistics, English curriculum and instruction, and EFL writing. She has been working and researching in English language teaching and learning, and has been involved in national English curriculum development for senior secondary vocational education and College English education.

Appendix 1: Composite analysis of diverse assessment elements

The figure below should be read as follows: Column 2 contains the analysis of the amalgamated five datasets of 158 items plus the anchor items. Column 3 contains the 53-item LTE test, Column 4 the 65-item CET test, Column 5 the 22 CSE-referenced self-assessment ratings, and Column 6 the 16 CEFR-referenced self-assessment ratings.

1	2	3	4	5	6
Test takers	Composite analysis (including *anchor test*)	LTE test	CET test	CSE Can Dos	CEFR Can Dos
					
Test takers	Composite dataset	LTE test	CET test	CSE Can Dos	CEFR Can Dos

CEFR JOURNAL—RESEARCH AND PRACTICE
VOLUME 5

Development of a Framework of Reference for Sign Languages and Reference Level Descriptors for Czech Sign Language

Martina Hulešová, Charles University

https://doi.org/10.37546/JALTSIG.CEFR5-4

The CEFR is a well-known, useful, and widely exploited tool used for many European languages, as well as in some non-European language contexts. Most of the contexts where the CEFR is used concern spoken languages. However, regarding sign languages, there have only been a few attempts to explore how the CEFR might be adapted and modified. In 2019, a CEFR-related project started in the Czech Republic (as one of five key activities of a bigger project – called APIV A – that concerned inclusion of users of first languages other than Czech) with the original aim of adapting the outcomes of the ProSign project[1] led by The European Centre for Modern Languages, which basically meant creating proficiency scales for the description of Czech Sign Language. However, it became clear that a mere translation or a slight adaptation is neither possible nor appropriate. Therefore, the project became much broader, and, in the end, two original comprehensive and interrelated documents were developed: a general Framework of Reference for Sign Languages and a more specific Reference Level Descriptors of Czech Sign Language. Both documents are bilingual: in written Czech and translated into Czech Sign Language.

Three main topics are discussed in this article. Firstly, the content and the processes by which these two documents were planned and published are described. The rationale for their development is presented, and the approaches, including blind alleys, doubts and their solutions discussed. Secondly, challenges faced by the writing team are presented, for instance the collaboration of Deaf and hearing colleagues, the collaboration of hearing linguists with Czech Sign Language teachers with no linguistic background, terminological issues, given that sign languages in general, and the Czech sign language specifically, are so-called less-taught languages. Finally, problems and challenges related to the features of the Czech deaf community, such as the specificity of culture, language modalities, the absence of standardisation, research, and the lack of recognition of the language as a fully-fledged code, are presented.

Key words: sign language(s), framework of reference, reference level descriptors, deaf community

1 Introduction

The CEFR and the CEFR Companion Volume are useful and well-known tools used in many European contexts. They have also been adapted for some non-European contexts where the CEFR impacted local language policies (Canada, Thailand, Malaysia, Uzbekistan) or the way languages were described, assessed or taught (Japanese, Arabic), as described, among others, by Salwa (2021, Savski (2020), Soliman (2017) or Khatamova (2018). However, the implementation of the CEFR has mostly been concerned spoken languages, i.e., audio-oral languages. For sign languages (SL), working within the visual-manual environment, there have been only a few attempts to explore how the CEFR might be adapted and modified for learners and users of SL, for instance, in a project at the Zurich University of Applied Sciences (ZHAW), funded by the Swiss National Research Programme, where descriptors and scales for SL competence complemented the CEFR scales, or, in the ProSign Project, led by the European Centre for Modern Languages, where

1. When the ProSign project is mentioned, we refer to the first ProSign project realised between 2012 and 2015.

variants of some of the proficiency scales and descriptors were developed for the use in SL contexts. The descriptors and scales for SL from the ProSign project were also used in the CEFR Companion Volume (CV) (Council of Europe [COE] 2020: 49), together with a specific chapter on signing competencies describing aspects of competencies unique to SL. These proficiency scales are accompanied by texts explaining their theoretical background and key concepts operationalised in them.

In 2017, a project called APIV A started in the Czech Republic to support the inclusion of language communities with different L1 languages, which would include the community of users of the Czech Sign Language (CSL). At the time of preparing the project documentation and applying for European funding for APIV A, only the ProSign project outcomes were known to the project promoters, and neither in the later stages information about similar projects was encountered. Also, the CEFR CV existed as a provisional version. Consequently, the CEFR 2001 and the ProSign documents were the main sources of information. Originally, it was envisaged that a translation of the ProSign documents and its use as a framework for the description of CSL competence would be the only primary outcome of the APIV A project and that it would be sufficient with some additional texts. Although the ECML project with the ProSign document as its most salient outcome was an important step on the way towards the wider visibility of SL, it turned out not to be comprehensive and consistent enough for the purpose of a thorough description of a particular SL. This was largely because complex information describing the specific modality of SL and the contrasting nature of the sign and spoken language was missing. Therefore, when work on the project relating to CSL began in 2019, the original intention to take the ProSign outcomes and to translate and slightly adapt them for the description of CSL was perceived as insufficient. The project became much broader. It was decided to create a coherent descriptive framework for SLs and CSL specifically. For this, it was necessary to have a theoretical foundation as well as a clear idea about the resulting framework's potential practical applications. Both ambitions would require a significant amount of theoretical research, consultancies, discussions, and cross-language collaboration, as well as a community willing to take part in this initiative.

In the following sections, the most important areas related to the development of both frameworks are described. Section 2 provides a brief overview of the project, the development, the content, the purpose, and the target group of readers of the two main outcomes (Framework of Reference for Sign Languages and Reference Level Descriptors for Czech Sign Language) and their relationship to the source documents (CEFR and ProSign). Section 3 describes the methodology, the workflow and some of the challenges, such as the need to work in two different language modalities and the problems related to the translation. Section 4 focuses on the approach to the validation, and explains how the validation was carried out and in what aspect was specific. Section 5 reflects the lessons the team learnt during the project. The last Section 6 ends with some concluding remarks on the benefits the project outcomes might bring for the Deaf community, Czech signing linguists and the Czech Deaf community in general.

2 An overview of the documents

Despite initial challenges, a late start, and issues related to bilingual-bimodal teamwork (which required constant linguistic and transcultural mediation), two comprehensive and interrelated documents were eventually developed: a *Framework of Reference for Sign Languages (FRSL)* and a comprehensive *Reference Level Descriptors for Czech Sign Language* (RLDCSL)[2]. The FRSL is a reference document that is intended to be read by a broader public, especially teachers, and students of SL, authors of syllabi, and curricula. The RLDCSL is a descriptive tool that collates notions and concepts specific for CSL. As the ProSign document was insufficient for their purposes, the authors went back to the original sources, i.e., to the CEFR 2001 and later to the CEFR CV 2020. As the CEFR CV writing team argues, "[m]any other CEFR descriptors are actually applicable to SL since it is used to fulfil the same communicative functions" (COE 2020: 49). In case of FRSL and RLDCSL, not all chapters, illustrative scales and descriptors from the CEFR

2. The project APIV A ended in November 2022.

were modified and used, on the one hand, and on the other hand, some other scales not included in the CEFR were developed: r*eception of artistic or entertaining texts in visual media or live; production of longer factual/expository texts; control of phonetic and phonological aspects of a SL*. The authors decided to use from the CEFR CV only very few texts and scales (roughly estimated at less than 10%) that were considered as the most relevant for the purpose, target readers and the SL and CSL. Several originally developed texts and scales were added by the Czech team as they were felt to be missing in the CEFR CV but important. In conclusion, the main source of inspiration was the structure of the CEFR and the interaction between the content of both CEFR for spoken languages and the specific nature and needs of the SL.

The resulting documents, FRSL and RLDCSL, contain theoretical chapters which explain important background information, the rationale for the content included and the approach adopted, as well as introducing the proficiency scales for SL (FRSL) or Czech SL (RLDCSL). They share structural and content features with the CEFR, but some chapters, particularly in terms of their level of specificity, are different. This was caused by the need to respect the nature of the SL in terms of their linguistic structure and the extent of knowledge (e.g., regarding the linguistics, applied linguistics and didactics) of potential users, especially within the Deaf community. In comparison with the CEFR, the FRSL and the RLDCSL use a more explanatory approach, trying to balance the specific terminology with the commonly used language in both systems, i.e., the spoken and the SL. Whilst the FRSL is intended to be read as a reference document by a broad but informed public (e.g., teachers, lecturers, students of SL, authors of syllabi, curricula, teaching materials, and assessment tools), the aim of the RLDCSL is to be the first comprehensive collation of the notions and concepts specific to CSL. In addition, it is intended as a tool that might be used for preparing language course syllabi and in the creation of teaching and assessment materials. In short, it is intended specifically for those involved in teaching and learning CSL.

2.1 Framework of reference for sign languages

The FRSL is intended as a reference document for SL in general and therefore, it is supposed to be read before the RLDCSL. Chapter 1 introduces the target group of readers, explains the history and the development of the documents that were sources for the CEFR. Chapter 2 describes how the sources were adapted when developing the FRSL. Chapter 3 explains the key concepts (e.g., reference level, descriptors and scale), and how they relate to each other, and other concepts mentioned in the following chapters. The text continues explaining in detail the role of the communicative situations, and introduces the notions of communicative spheres, communicative factors and communicative activities and strategies. Chapter 4, in which communicative activities and strategies are explained, is the key part of the FRSL. It contains and then defines proficiency scales from Pre-A1 to C2 and describes what the user of SL can do in and with SL in different communicative situations. Chapter 5 describes the language competencies (or signing competencies, as stated in the CEFR CV) of the SL users: linguistic competence (phonetic-phonological competence, grammatical competence, lexical competence), pragmalinguistic competence and sociolinguistic competence.

2.2 Reference level descriptors for Czech Sign Language

Although the FRSL offers a general description of the language behaviour of the SL user in different communicative situations (defined by the proficiency scales), the RLDCLS introduces knowledge, topics, and skills whose acquisition enables the CSL user to perform language activities described in the FRSL. Only levels Pre-A1–B2 are included in RLDCSL, which is in line with the project brief. However, there are also pragmatic reasons why C levels are not included. There are practically no signers at C1 or C2 levels among the users of CSL as a non-L1 language; there are no teaching materials, corpora, and very few students interested in continuing studying in courses at these levels. As a consequence, examples of language production at C levels were not available.

Chapter 1 of the RLDCSL introduces the material itself, its content, structure, how the text relates to the CEFR, it explains how the text should be read and used. And levels A1-B2 are characterised briefly in this chapter. *Chapter 2* provides basic information about the nature, modality, and key terms of CSL, such as the manual and non-manual language means simultaneity, linearity, and iconicity. *Chapter 3* describes the essential language structures and rules, and it also contains a list of grammar structures and lexical units related to topics and reference levels A1-B2. *Chapter 4* summarises sociocultural knowledge and skills related to the community of Czech SL users and defines what is expected (regarding sociocultural knowledge and skills) at levels A1-B2. *Chapter* 5 describes language knowledge and skills at A1-B2 levels from the perspective of 14 thematic areas and the most common expressions related to these areas are listed in the subchapter called *Vocabulary*. In the Czech language version of the documents, these expressions are represented by dictionary entries that are the equivalents of the signs. Each chapter has slightly different content and graphical structure that respect the content and approach felt as the most effective with regards to the target users.

Both documents also contain examples of the language in use; the FRSL attempts to exemplify the general principles of SL, and the RLDCSL contains examples of the language use, which is specific to CSL. Both contain scales with illustrative descriptors with Can Do statements. The reference scales in the FRSL go from Pre-A1 to C2 as their intention is to describe the features common to SL in general, whereas the RLDCSL contains scales from A1 to B2.

2.3 Glossary

The Glossary accompanying the FRSL contains terms that were considered key or important ones, as well as those which were found challenging during the validation process both by internal hearing and not-hearing colleagues, as well as by the external reviewers. The explanations attempt to be in line with the current thinking in Czech linguistics. In a broader sense, the FRSL and RCSL may serve as a source of key terms and the metalanguage for all those involved in the (C)SL community, as their descriptive language mirrors and follows important current trends in Czech linguistics.

Both the FRSL and RCSL and the Glossary exist in two language versions: in Czech and CSL. The website[3], in addition to other information, will contain both language versions and a downloadable interactive PDF version in Czech.

2.4 The purpose of the documents and the target group of users

The initial definition of the target group for the FRSL and the RLDCSL in the Project Chart was very broad and allowed for many interpretations[4]. On the one hand, it allowed for a later broadening of the project's scope to the needs that emerged as the team gained knowledge, for instance, having two separate documents with differentiated purpose and content (FRSL and RLDCSL). On the other hand, it was challenging to define the target user and, therefore, to determine the scope of the documents, the level of detail, and the language used. Therefore, one of the main discussions at the beginning of the project work focused on defining the target group of users. After a series of discussions within the team and within the broader community of signers and linguists, the main target group of users was defined as the teachers of CSL as a foreign or second language, i.e., mainly deaf teachers in courses of CSL for hearing students.

Parallel to the discussion about the target user of the FRSL and the RLDCSL, a discussion about the articulation of the purpose of the documents took place. As had happened with the CEFR, which was taken as a prescriptive instead of a descriptive tool at the beginning of its existence, a similar reaction occurred when preliminary versions of the FRSL and RLDCSL were presented to members of the Czech

3. The website was launched at the end of 2022: https://cefr-czj.npi.cz/
4. The intended uses originally went from FRSL being a framework for developing exams for SL interpreters to a resource for teachers of deaf students at primary schools.

deaf community, as well as to other stakeholders with limited knowledge of or experience with the CEFR[5]. The clarity of the purpose and the target reader were also two of the areas investigated in the validation process (see section 3.6), which brought important feedback to the writing team. The validators (validating the RLDCSL) and reviewers (validating the FRSL) pointed out the need to describe more clearly and explicitly the intended purpose and to precisely define the target users of the documents. To this end, separate chapters on how to read and use the FRSL and the RCSL were added.

3 Methodology: Teams, workflow, and challenges

External and internal members worked together as part of a broader team. As the teams and their members had different knowledge and specialisations, each team had different duties, such as writing texts and scales, searching the literature, providing consultancies, leading validation, translating, and editing. The relation among the teams is outlined in the scheme in Figure 1. The central groups are represented by the central five subgroups connected with arrows that represent the directions of communication. The *Search team* and the *Website team* worked more at the beginning or in the later stages of the project, respectively. External colleagues took part in specific points of the project.

3.1 Teams and their duties

The *Linguistic team* was the main writing team. It was composed of Czech signers and Czech speakers, most of them linguists. The *Linguistic team* and the *CEFR team* asked for support from the *Research team*. They looked for relevant literature, mainly at the beginning but also during other project stages. The *CEFR team* had two roles: first, they provided consultancy, information, familiarization, and initial training in working with the CEFR and related materials. Later, they collaborated closely with the *Linguistic team* as co-authors of some chapters. They supported the *Linguistic team* with their expertise in the CEFR, provided feedback and advice on issues relating to the adaptation of the CEFR to the new SL and CSL needs, as well as on issues concerning the alignment of the new scales to the CEFR.

The *Validation team* was created in one of the later stages of the project. They were responsible for gathering the feedback on the texts, leading the validation of the documents, and providing feedback on the created documents. The *Validation team* collaborated with both internal and external colleagues. As these were Czech speakers but also Czech signers, close collaboration with the *Translation team* was needed.

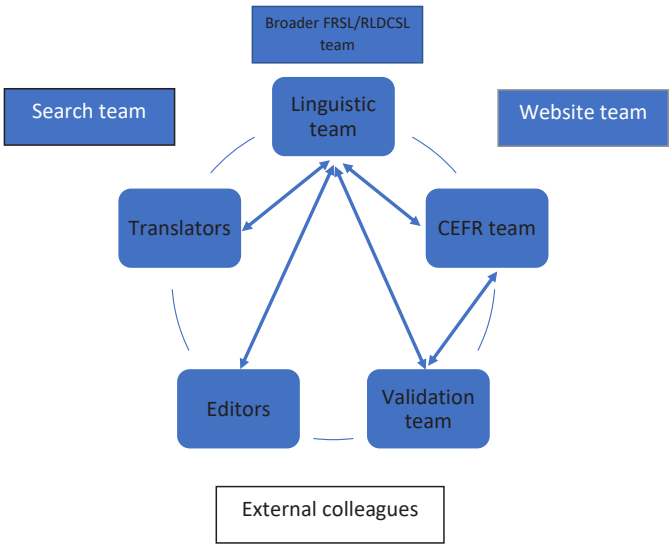
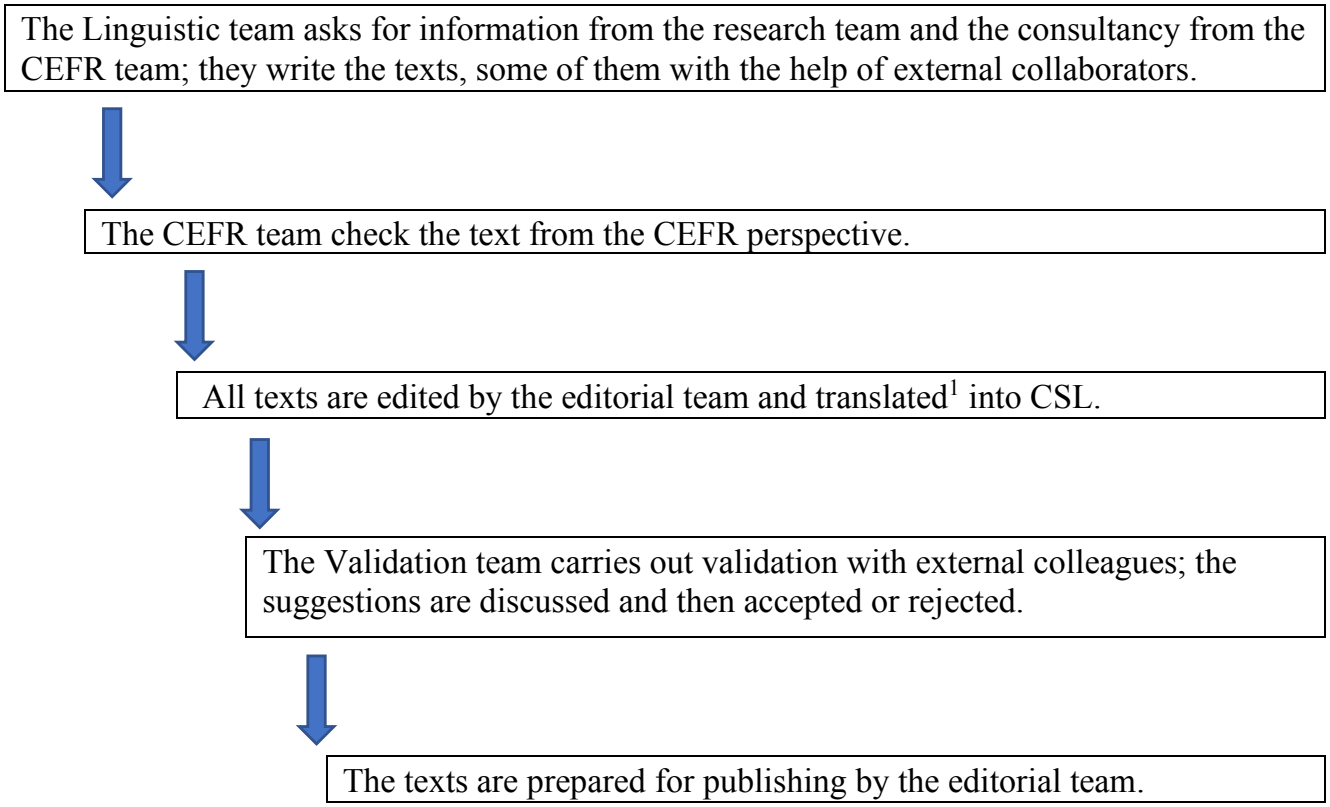


Figure 1. Teams and their collaboration

5. During the project, there were about 10 dissemination panels where the results were presented and discussed with the public. Also, regular reports to the Ministry of Education, Youth and Sport had to be presented.

3.2 Workflow challenges

The initial workflow was envisaged as follows:



However, in reality, the workflow was not linear and smooth. Writing the text required much more involvement from the CEFR team and the external collaborators. In addition, the validation required much more involvement from the Linguistic team and the Translation team. The Translation team had to retranslate many texts several times. Some parts of the texts were translated again based on the validation findings.

After several cycles of feedback, discussion, clarification and redrafting, the text was given to the Editorial team, who then prepared it for the final translation. Once the translations had been completed, the text was read and reread in consultation with the Linguistic team and sometimes also with the CEFR team, and in case it was needed, redrafted, and translated again.

3.3 Living and working in two language modalities

Probably the most basic and salient difference between spoken and SL is the modality, i.e., the way the languages exist and how they are used in and for communication. Spoken languages are of an audio-oral nature, the physical perception happens through hearing, the production by the speech organs and the meaning lies in the sounds and the context. SL is of a visual-motor nature, i.e., the message is perceived by sight, produced by the body parts, movement and space, and the meaning is created by the shape, position and movement of the hands, body, head, and face muscles, and, as with spoken languages, in combination with the context. The different modality of SL brings about other challenges. For instance, SL do not have a written form and the message can only be live or video recorded. As language is closely related to thinking and culture in the broad sense, languages with different modalities also assume different cultural, sociocultural as well as interpersonal contexts, and behaviours. Norms, relations, and taboos are not always shared or equally understood.

From the simplified explanation above, it is evident that the hearing and signing colleagues did

not share the same communicative space, and they were not able to communicate directly without translation or mediation. Usually, for less formal meetings, hearing members of the linguistic team interpreted the communication; for more formal meetings or broader team discussions, interpreters were required. All of this had an impact on the organisation and operativity of the communication. Furthermore, a group of translators was needed to disseminate information to the public, and the mediator had an important role during the long-term validation process, as they helped beyond just transferring the meaning. In fact, they also carried out a transcultural transfer to avoid socio cultural misunderstandings or embarrassing situations.

3.4 The languages of the documents and the translation

One of the frequent questions the people outside the project asked the writing team why the documents were written first in Czech and only then translated into CSL. It is a perfectly legitimate question for many reasons: the documents are about SL or CSL, the most important target group are primarily people with CSL as their L1, and due to the different modality of SL, the thinking, living and world-perception are shaped differently for CSL and, thus, generally less accessible to Czech speaking linguists or experts.

The honest answer to the question of why the team decided to work in the spoken Czech first is quite simple: it was difficult to find enough colleagues within the Czech Deaf community that would meet the criteria for being able to work on the project (for instance, to be familiar enough with the CEFR and its use, to be able to read and to know the terminology in English and Czech; to have a background in linguistics and didactics of foreign languages). It should also be mentioned that the CSL is in an early stage of development in some political, educational, and especially legal aspects. The project itself was proposed and written by both Czech signers and Czech speakers who had contacts with the Czech deaf community and who were sensitive to the need to fill the existing gaps in terms of the linguistic and didactic tools for describing CSL. Subsequently, Czech speakers and Czech signers agreed on the need to have this tool, but when the project leaders looked for team members that would meet the requirements and were willing to take part in the project, it was much more challenging to find Czech signers than it was to find Czech speakers[6]. Due to the limited availability of educational programmes designed specifically for members of the Czech Deaf community, there is a relatively low number of Czech signers with a university degree in linguistics. As a result, the team's make-up favoured Czech speakers with more experience in linguistics than Czech signers. Therefore, it was decided to work on the documents in spoken Czech first, to distribute the tasks according to the profiles of each member of the project teams and translate the texts later into the Czech signed language. This decision required a lot of coordination across the team, constant monitoring of the workflow, and very intensive communication among the teams.

The decision also led to some negative consequences for the timetable, especially for editing and translating. These had to be postponed as much as possible as we wanted to translate as complete a version of the documents as possible, and some parts were re-translated several times. This was caused by issues with a non-standardised translation that resulted in misunderstandings by Czech signers and by the need to implement the findings emerging from the validation.

3.5 Working with and in a non-standardised language

The CSL lacks a standard form. Thus, it might be characterised as a language with a highly individual but particularly lexical variability (Hynková Dingová 2020). This variability is conditioned geographically and socially. Although CSL has become more frequent in the public space, which has had a positive impact on the standardisation (as the language used in the public media is taken as the model of

6. All the Czech speakers in the linguistics team were highly proficient in Czech Sign Language, but not the members of the CEFR or the validation team.

use), the variability is still very high. The lack of a language standard meant that the translators and the linguistic team members struggled due to the lack of equivalents or standardised signs for many concepts (relating to the CEFR, linguistics, abstract ideas, and metalanguage) in the non-standardised CSL. Unfortunately, this was discovered quite late, after the first piloting of the validation method during one of the dissemination panel meetings. Only after this feedback, problems with misunderstanding the texts were discovered, and issues in translations were detected. The team realised that coherent, acceptable equivalents for the terms from the CEFR CV (originally used in English) were not used consistently in CSL. For some concepts, their equivalents were even missing in CSL. It was challenging to find agreement on the translation form for many expressions, as the standard equivalent did not exist and had to be created and agreed upon among the translators and the community.

4 Validation

The initial intention to follow the approach towards the validation as described by North (2007, 2020), North and Piccardo (2019) and CEFR CV (2020), both with panels of Czech signers and Czech speakers, turned out not to be feasible after the piloting phase that emulated the procedures described in the above-mentioned literature. The main conclusions from the piloting that led to the decision to change the validation approach were (a) the modality of the SL that makes the process difficult for presenting the activities with scales and descriptors, (b) the level of preparedness of the deaf colleagues and their lack of experience with similar activities; (c) the sociocultural aspects specific to the Czech Deaf community where the members are not used to be trained or taught by a hearing person from outside the community; (d) given that working with descriptors and scales presupposes a certain level of knowledge of the descriptive scheme itself, its language and metalanguage, the terminology; in fact, for the Czech Deaf community, the texts were incomprehensible without a mediator (usually from the Linguistic or the CEFR team); (e) very few members of the deaf community had experience with workshop-based activities. Therefore, the validity had to be re-conceptualised and a new approach towards validation had to be applied, different both in terms of the content (WHAT was validated) and in terms of methods (how the content was validated). The validity of the documents was conceptualised as the agreement between three facets: the theoretically defined purpose of the documents based on the needs of the main target group of users described theoretically, the validators and reviewers representing the target group of users representing, and the content of the FRSL and RLDCSL that describe the construct of SL (or CSL) from different perspectives, such as learning, teaching and assessment. The validation emphasised the aspects of comprehensibility for the users, completeness, usefulness, and balance between the level of expertise and accessibility. One of the most important indices of the validity was the attitude of reviewers and validators towards the documents, specifically, how the attitude developed during the validation process and what impact the validation had not only on the participants but also on the broader community.

Basically, the FRSL was validated by *reviewers*, i.e., a mixed group of three hearing and deaf experts (in linguistics, pedagogy and the CEFR). They were asked to provide a detailed structured review of the whole document except the scales. It was expected they focus on four main areas: comprehensibility, completeness, usefulness and balance between expertise and accessibility. The questions they were asked were not too specific and left room for broader answers and deeper thinking about the areas. Reviewers were given the whole document and were asked to provide a provisional review first, then to meet with the validation teams for consultancies, and only after this write the final review. They worked independently, always having the possibility of consultancies with the CEFR teams with whom they consulted several times. A final meeting was organised with two aims: first, to give experts the opportunity to ask for explanations and to see the changes made in the FRSL after the validation, and, second, to give an opportunity to the CEFR team to explain some issues that proved to be unclear to the experts.

The RLDCSL was validated by *validators*, i.e., a group of Czech signers with profiles reflecting the target group of RLDCSL users. The outcome was a written or video-recorded review. The key person was the *mediator*. The mediator was a signing colleague hired to lead the group of signing validators and mediate the content and the processes of validation, which had been designed by the hearing members of the validation team. He was trained by hearing colleagues who provided him with support, consultancy, and explanations. This was done with the help of interpreters.

The RLDCSL validation had several rounds. Each chapter was validated separately by a slightly different group of the validators against a set of topics-questions related to these chapters. These questions were mediated by the mediator because validators and the mediator were exclusively deaf colleagues. However, the materials were in a written form prepared by Czech speakers from the Validation team.

Validators focused basically on the same aspects as reviewers (comprehensibility, completeness, usefulness and balance between the expertise and the accessibility), but the questions were formulated in a completely different way. They were introduced by a short explanatory text describing the chapter in question, they were worded very specifically and explicitly, and there were several questions targeting the same aspect. Validators were also asked to rely on their experience as teachers of the CSL.

The mediator and validators met several times at individual consultancy meetings. These meetings were not translated simultaneously. They were transcribed for the purpose of reference and the Validation team. After each individual meeting, the mediator met with the Validation team for consultancy and to plan the next meeting. The final group meetings of each group of validators with the mediator were recorded. The aim of these final meetings was to receive answers to the questions prepared by the Validation team. These final meetings were interpreted simultaneously because the Validation team took part in the discussions.

5 Lessons learnt

When looking back, all parties involved in this project learnt several important lessons. Firstly, in similar projects in the future, the writing team should communicate more closely across the teams and prevent misunderstanding and the creation of different levels of shared knowledge. Secondly, areas of responsibility should be better defined, distributed and incomplete tasks should be discussed immediately at regular meetings of the project members. As well as being regularly organised, these meetings should be attended by members from across different teams. Thirdly, more training (especially at the beginning of each stage) in topics related to the project aims, should be provided, and terminological and conceptual issues should be discussed. Finally, in addition to the training, a shared database of frequent questions and terms and concepts should be created. These should be bilingual and accessible to all team members. On the other hand, the project also had a very positive impact. The Czech deaf community and Czech-speaking experts collaborated closely on such a big project for the first time, and we hope this prepared the floor for future collaboration of both parties and also for the emancipation and more independent work of deaf colleagues and their major involvement in similar projects. The initiative for future projects in CSL should come from the Czech deaf community, and the project should be led by deaf colleagues.

6 Concluding reflections

As it was mentioned before, both documents complement each other. The FRSL introduces the Czech Deaf community to a theoretical description that might be conceptually challenging. However, when considered in the light of the RLDCSL, it can be understood thanks to the specific examples the RLDCSL provides.

The project itself, the outcomes (the FRSL and the RLDCSL), has had a positive impact on the Czech deaf community as a whole. Their members were invited to take part at different stages of the project

(as consultants and validators), and they were also invited to the dissemination panel meetings, which were organised during the three years the project ran. They became aware of what the project was about, and they were informed about the goals, outcomes and the use and usefulness of the project.

Both the *Framework of Reference for Sign Languages* and the *Reference Level Descriptors for Czech Sign Language* are highly significant for the Czech deaf community. They provide a unique theoretical background and description of CSL in relation to the CEFR levels, as are common across other European languages. This will make communication about the use, teaching, learning and assessment of CSL possible within and outside the Czech Deaf community. Potential users of the FRSL and the RLDCSL will have a descriptive tool that enables them to understand each other when communicating in their areas of interest, in teaching, preparing syllabi, planning curricula, assessing, or learning CSL. It might also be a positive step towards standardisation, at least in pedagogy, teaching and learning CSL as a foreign or second language. The descriptive and illustrative nature of their content might increase the comparability of the courses, materials, and assessment approaches and thus improve the mobility of students and teachers of CSL. They might also help improve the quality of teaching and assessment of CSL, to support the production of course materials and content, and to give teachers a common language in which they can communicate about CSL and its users, and open a broader discussion about the SL itself, its nature, the ways to teach and learn it effectively, and thus support CSL on its way to become a fully-fledged language.

7 References

Council of Europe (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion Volume*. Council of Europe: Strasbourg, https://www.coe.int/lang-cefr (accessed 30 January 2021).

Hynková Dingová, Naďa. 2020. Situace českého znakového jazyka: významný předpoklad pro uchopení tlumočení v sociálních službách. *Sociální služby. [The Situation of the Czech sign language: significant assumptions for interpreting in social care.],* XXII (červen-červenec). 42-45. https://ujkn.ff.cuni.cz/wp-content/uploads/sites/64/2020/08/Situace-%C4%8Desk%C3%A9ho-znakov%C3%A9ho-jazyka.pdf (accessed 2 May 2022).

Khatamova, Salomat. (2018). Impact of the CEFR on educational policies. Иностранные языки в Узбекистане, [Foreign languages in Uzbekistan], (2), 138-141.

Leeson, Lorraine, Beppie van den Bogaerde, Christian Rathmann & Haug Tobias. (2016). *Sign languages and the Common European Framework of Reference for Languages Common Reference Level Descriptors.* Council of Europe: Strasbourg. https://www.ecml.at/ECML-Programme/Programme2012-2015/ProSign/tabid/1752/Default.aspx (accessed 25 April 2020).

North, Brian. (2020). CEFR Validation Reference Paper – with appendices. https://www.researchgate.net/publication/350156412_North_CEFR_Validation_Reference_Paper_-_with_appendices (accessed 31 May 2022).

North, Brian. (2007). The CEFR illustrative descriptor scales. *The Modern Language Journal* 91. 656-659. https://doi.org/10.1111/j.1540-4781.2007.00627_3.x.

North, Brian & Enrica Piccardo. (2019). Developing new CEFR descriptor scales and expanding the existing ones: constructs, approaches and methodologies. *Zeitschrift für Fremdsprachenforschung,* (30/2), 142-160. https://www.academia.edu/41330365/North_B_and_Piccardo_E_2019_Developing_new_CEFR_descriptor_scales_and_expanding_the_existing_ones_constructs_approaches_and_methodologies (accessed 30 April 2022).

Salwa, Mohamed. (2021). The development of an Arabic curriculum framework based on a compilation of salient features from CEFR level descriptors. *The Language Learning Journal*, https://doi.org/10.1080/09571736.2021.1923781.

Savski, Kristof. (2020. Local problems and a global solution: examining the recontextualization of CEFR in Thai and Malaysian language policies. *Language Policy 19*. 527–547. https://doi.org/10.1007/s10993-019-09539-8.

Soliman, Rasha. (2017). The Implementation of the Common European Framework of Reference for the Teaching and Learning of Arabic as a Second Language. In Wahba, K.M., England, L., & Taha, Z.A. (Eds.). *Handbook for Arabic Language Teaching Professionals in the 21st Century (1st ed.).* Routledge. https://doi.org/10.4324/9781315676111.

8 Biography

Martina Hulešová graduated in Spanish philology at Charles University (Czech Republic), later obtained her MA in language testing at Lancaster University (UK) and PhD in didactics of foreign languages at Masaryk University, Czech Republic. Her research interests include standard setting, validation, training in language testing and assessment. She works as quality manager for the Research and Test Centre at the Institute for Language and Preparatory Studies at Charles University and as a freelancer she provides consultancy for other projects related to language assessment, Czech sign language, diagnostic tests and others.

CEFR JOURNAL—RESEARCH AND PRACTICE
VOLUME 5

Implementing a localized version of the CEFR-based curriculum in Israel

Tziona Levi, Israeli Ministry of Education
Simone Duval, Israeli Ministry of Education

https://doi.org/10.37546/JALTSIG.CEFR5-5

This research addresses the implementation of the new Israeli English curriculum based on the CEFR and its introduction to lead-teachers, teacher-mentors and teacher-educators defined as 'expert' according to Israel's Professional Standards Document (Ministry of Education 2019). To create reciprocal study and interaction with the curriculum document while disseminating it to EFL teachers throughout the country, these experts were led to conduct a multilevel interactive discourse, characterized by the ripple effect metaphor within Professional Learning Communities (henceforth PLCs). Examination of this interactive discourse within the PLC framework reflects the incorporation of a unique application design that draws on EFL curriculum implementation as a national policy and concurrently provides insight into the delivery of the curriculum designed to elicit critical meditative conversations. While applying a PLC setting, we demonstrate collaborative dialogues and knowledge construction by participant 'experts' as they learn the curriculum through social interaction, activating conceptual curriculum language as a mediational tool and verbalizing or 'languaging' the meaning making process (Swain 2006; Watanabe and Swain 2007). Thus, we ascertain that the expert-teachers' knowledge of the curriculum is constructed within the PLC structure as they explore methods to mediate the curriculum. To capture the scope of the interaction and delineate this knowledge construction, we collected, transcribed and analyzed asynchronous logs written by each participant, and four collaboratively written (socially constructed) group logs which together form the source of the current qualitative study.

Keywords: CEFR, professional learning communities (PLC), social interaction, curriculum implementation, curriculum mediators

1 Introduction

Israel's education system has a centralized national English curriculum ensuring the uniformity and standardization of teaching objectives. K-12 English teachers are required to follow a structured framework, as reflected in high-stakes testing; what is not tested is not taught (Shohamy 2001). This testing culture affects how policy is interpreted and implemented by the teachers in class.

English is considered the first foreign language in Israel. However, being an extremely diverse society, language learning in school has different meanings for different groups. For example, Arabic-speaking students consider English as their third or fourth language as they learn spoken and then written Arabic, then Hebrew as the national language, and then, English.

Considering the centralized nature of the education system, implementation of the English curriculum was previously imposed as a top-down policy by the Ministry of Education that determined content and process to ensure fulfillment of goals in the transition to practice. Teachers' perceptions of the changes were not considered, and this often resulted in resistance to the changes.

The introduction of the CEFR-aligned curriculum into the Israeli context required a definite change in teachers' mindsets and required a deliberate long-term and multi-stage implementation plan.

The most obvious change was the use of CEFR terminology. Previously, teachers prepared or used materials organized according to general benchmarks. The CEFR curriculum relates to task-based activities in the form of Can Do descriptors offering a more student-centered approach and relating to learners as language users who must develop a sense of responsibility in developing their language skills. Additionally, as language users, students are encouraged to focus more on productive skills.

Furthermore, from the earliest stages of implementation, it was clear that the Backward Design Framework to lesson and unit planning was best suited to the new CEFR-aligned curriculum. This can be seen as another change from the previous curriculum, which did not singularly focus on one specific unit or lesson planning framework. Backward Design requires unit rather than lesson planning, and a need to decide on the desired outcome to allow tying teaching, learning and assessment into one cycle.

Other changes were the inclusion of mediation and interaction as language activities, and preparation of a list of most frequently used vocabulary (BANDS 1-4) according to the levels of progressions as outlined in the CEFR. Consequently, the learning of the curriculum required a less top-down process. Teachers were encouraged to be involved in the curriculum implementation stages and provide input and reflection while internalizing the new concepts. In addition, the construction, writing and implementation of the CEFR-aligned Israeli English curriculum document was different.

Once the draft document was published, the English Inspectorate began the task of turning policy into practice. Relaying the document as a draft was intentional and reflected a belief that the various stages of implementation required the direct input of active teachers to receive ongoing feedback on the appropriateness of the document for Israeli classrooms, suggesting a combined bottom-up and top-down approach to implementation. Thus, the first stage of dissemination required a professional development (PD) course designed to present the curriculum document to English lead-teachers and teacher-mentors. The purpose of the first eight-session course and the following 40 courses of 30 hours each was to allow the lead-teachers to develop a deep understanding of the rationale during the writing of the curriculum. During this stage, the lead-teachers and teacher-mentors in the first PD courses asked questions, which obliged the writing committee to critically evaluate and review the written document through the lens of the course participants and other teachers. Some of the challenges experienced were resistance to the changes, a lack of understanding of the new terminology, and a clash with terms from the previous English curriculum in relation to newly introduced ones. Many of the issues raised by the course participants and teachers led to some revisions to the document, decisions for further implementation and a clearer perception of the practical implications of the policy document.

2 Context of study

The first draft of the A1 and A2 levels of the CEFR-adapted curriculum was published for Israeli schools in 2020, and the introductory courses were followed by the start of the second stage of implementation. This second stage focused not on the 'what' of the new curriculum but rather on the 'how' of its application. The course was designed for experienced or influential EFL educators from schools around the country to allow for curriculum cascading and dissemination from them to other English teachers. The second stage aimed to focus more on how a smaller group of lead-teachers and teacher-mentors could develop their strategies to mediate the written document to teachers in the field.

At this time, many teachers had not yet completed a curriculum course or had done so but were still unable to adapt their teaching practices to the CEFR mindset. The English Inspectorate decided that the PLC framework was the most suitable and effective way to move to the second stage of implementation. The structure of PLCs encourages participants to collaborate as a community to consider their individual and collective roles within this process. These PD courses were aligned with the bottom-up approach of professional learning (DuFour 2004).

Based on the success and some comments of participants in the first course, such as *"Teacher colleges have to make sure their courses – especially methodology and practicum are aligned with the principles in*

the curriculum" (G1, T7) and *"For preservice teacher training, this means that those devising program content and courses need to be aware of the principles and rationale of Curriculum 2020, and teach accordingly"* (G1, T2), it was decided that the second and third courses of stage two would not just be for EFL lead-teachers and teacher-mentors but should also include college preservice teacher educators. Those chosen to participate were expected to have attended the introductory curriculum course or have a good understanding of the curriculum document. Teacher training instructors in colleges were invited to join the second and third round of these courses so that they could learn alongside teachers in the field. Thus, participants were encouraged to relate to the curriculum as a continuum from schools to colleges so as to create continuity and allow for long-term planning starting at the preservice level. 81 participants in total took part in three courses over three years. EFL school lead-teachers, teacher-mentors and preservice teacher educators were recommended and chosen by their English inspectors and college heads to take part in these three PLCs as 'expert teachers,' as defined in the Israeli Professional Standards Document (Ministry of Education 2019).

2.1 Sociocultural perspective of the study

This study is theoretically informed by Vygotsky's theory of cognitive development (1978) known as sociocultural theory (SCT). Although proposed to investigate children's cognitive development, SCT has been shown to be relevant to L2 and FL (Foreign Language) teaching and learning (Frawley and Lantolf 1985; Lantolf and Thorne 2006; Williams 2013; Gánem Gutiérrez 2008, Gánem Gutiérrez 2013; Swain and Lapkin 1998, Swain and Lapkin 2000), which is the context of the participating lead-teachers and teacher-mentors described in this study.

Why is it useful to talk about curriculum implementation applying a PLC format from a Vygotskian theoretical perspective? How might Vygotsky's theory be helpful in developing the understanding of best practices in disseminating a curriculum to a group of English teaching experts whose role is to mediate their learning to other English teachers? To answer these questions, it is essential to discuss the concept of mediation. Vygotsky (1978) saw language as a psychological tool, that is, a tool that mediates thinking. He asserted that the most important tool for developing and mediating thinking is language. Vygotsky saw language as a symbolic thinking tool through which we can explain the central concept of mediation. Mediation refers to how humans use actual or symbolic artifacts to assist both their physical and mental thinking activity while developing their understanding and concept-based knowledge of the world (Lantolf 2006: 69). The most important of these tools is language, as it is used to mediate mental activity and how the world is observed and understood. Language allows the exchange of information, talking and thinking about the present and connecting to events unrelated to the current time and space (Lantolf and Thorne 2006: 201-202). Beyond these, language is used as the tool that mediates higher mental processes and can be described as 'the process of making meaning and shaping knowledge and experience through language' (Swain 2006: 98). When language is used to mediate conceptualization and problem-solving, meaning-making, verbalization or languaging takes place.

Some L2 studies demonstrate how verbalization is considered a source of learning (Swain 2006; Watanabe and Swain 2007) and how L2 learners make use of language as they both develop their own thoughts and collaborate with others (e.g., Donato 1994; Lantolf and Appel 1994; Ohta 1995). When learners interact and collaborate to generate thoughts, they are constructing through interaction and applying a collaborative dialogue (Swain and Lapkin 1998), that is also used to explain development and learning (Swain and Watanabe 2013).

In the current study, construction of curriculum knowledge and languaging were designed to enable curriculum understanding manifested by group logs written by colleagues and peers.

2.2 Relaying the English curriculum through PLCs

The PLC can be defined as any group with a common interest in education (DuFour 2004). PLCs have been identified as an effective setting for sustaining learning and for developing teachers' motivation, skill, positive learning, organizational conditions and culture within an infrastructure of support (Stoll et al. 2006: 221). PLCs are a world-wide initiative with broad and varied foci such as "a deep sense of moral purpose, knowledge of the change process, capacity to develop relationships across diverse individuals and groups, fostering knowledge creation and sharing, and the ability to engage with others in coherence making amidst multiple innovations" (Fullan 2003: 7). PLC features include: shared values and vision, collective responsibility, reflective professional inquiry, collaboration and enhanced group and individual learning (Stoll et al. 2006). A rich body of research has investigated PLCs as professional teacher learning (e.g., Stoll et al. 2006; Fullan 2003), however, there is still a need to scrutinize the process of participants' knowledge construction within the PLC. This study will look at how EFL lead-teachers, teacher-mentors and preservice teacher educators construct knowledge during a PLC.

2.3 Social construction of knowledge in PLCs

Scardamalia and Bereiter (2003) define knowledge construction as, "the production and continual improvement of ideas of value to a community, through means that increase the likelihood that what the community accomplishes will be greater than the sum of individual contributions and part of broader cultural efforts" (1371). This definition aligns with two key features of PLCs, interactivity and collaboration. Knowledge construction is a collaborative effort that relies on the interaction of the members of the community through discourse or professional conversations (Kim and Wilkinson 2019; Lefstein et al. 2020). The interaction of the members of the community is essential to the learning processes (Popp and Goldman 2016) within the PLC.

Therefore, observing the interaction and knowledge construction within a PLC of EFL experts affords an additional layer to this research.

2.4 PLCs for Teachers of English as a Foreign Language

Considering the importance of EFL instruction in certain countries, there is surprisingly little research on EFL and PLCs (Pang 2019). Borg (2006) relates to English language teachers as having distinctive characteristics from other subject-matter teachers. Teachers of EFL are committed to the learning of the English language, not only the pedagogy of teaching. Consequently, it is necessary to investigate foreign language teachers separately as a distinct group in addition to the PD courses they participate in. In their research on subject-matter PLCs for English Language Arts teachers, Popp and Goldman (2016) suggest the importance of expanding studies exclusively on knowledge building in a subject-matter focused PLC.

The focus of this study on EFL lead-teachers, teacher-mentors and preservice teacher educators defined as experts in EFL, broadens our understanding of knowledge construction related to EFL teaching in a PLC setting.

The aim of lead-teachers and teacher-mentors is to make the practitioner's knowledge visible for the novice and experienced teacher and provide the means by which such knowledge can be understood and subsequently implemented in the classroom (Becher and Orland-Barak 2016).

Tillema and Orland-Barak (2006) researched collaborative knowledge construction in professional conversations. They determined that activity in context, particularly collaborative inquiry, contributes to knowledge construction. The mentors who participated in that study stated that engaging in professional conversations with their colleagues was a necessary component of teamwork. Therefore, activity and participation, supported by a framework of conversation contributed firstly, to the success of the PLC by building a sense of community, and secondly, to the construction of curriculum knowledge (Tillema and Orland-Barak 2006).

3. Research questions

For this study, two research questions were posed:

1. What are the perceptions and attitudes of the Israeli lead-teachers, teacher-mentors and preservice teacher educators regarding their ability to mediate the implementation of the CEFR-based curriculum?
2. How do EFL lead-teachers, teacher-mentors and preservice teacher educators construct their knowledge of the CEFR-based curriculum in a PLC?

4 Methodology

This longitudinal study conducted a qualitative in-depth case study analysis of three PLCs. Researching the process of knowledge construction using more than one data collection type allows for a more comprehensive understanding of the process (van Schaik et al. 2019). Two sources of rich qualitative data were collected through: 1) written personal logs; and 2) group logs, and were analyzed through content analysis in terms of themes and codes.

4.1 Participants

The participants were 81 lead teachers, teacher-mentors and teacher-training instructors who participated in three PLCs from 2020-2022 (see Table 1). They were chosen by the English Inspectorate with the aim of improving, developing and enhancing teachers' mediation skills with structured guidance and opportunities to build a specially designed knowledge base (Ambrosetti 2014; Feiman-Nemser 2001; Genç 2016). All 81 teachers were elementary, junior high and high school English teachers and or preservice teacher educators from around Israel.

In the first course, 30 lead-teachers and teacher-mentors participated. In the second course, 13 of the 24 participants (54%) were preservice teacher educators. This was the first time that lead-teachers, teacher-mentors and college instructors were brought together in a course initiated by the English inspectorate of the Ministry of Education. The third course also aimed for this balance and 15 of the 27 (56%) participants were preservice teacher educators (see Table 1). The belief of the English Inspectorate was that in order to promote the continuum from the school system to the colleges, the college teacher educators need to be partners throughout the process. The connections, discussions and artifacts that came out of this collaboration were informative and granted a better understanding of the process of curriculum implementation over four years.

Table 1. ***Number of participants in each group as lead teachers, teacher-mentors and teacher-training instructors***

3 PLCs	Teacher-training instructors	Lead-teachers and teacher-mentors
Group 1, 30 participants Group 2, 24 participants Group 3, 27 participants	Group 2, 13 participants Group 3, 15 participants	Group 1, 30 participants Group 2, 11 participants Group 3, 12 participants
81 total participants	28 total Teacher-training instructors	53 total lead teachers and teacher-mentors

4.2 Data collection, analysis and research ethics

Within these three PLCs, all the participants documented their learning in four individual logs (Appendix A) and then in groups of three or four, they wrote group logs (Appendix B) throughout the entire process.

The aim of the individual log was to encourage the participants to independently consider four aspects of curriculum implementation: (a) mediation, (b) backward design, (c) recorded lessons as representations of practice, and (d) simulations (live case studies that were acted out and debriefed, and then analyzed by the participants). After writing each individual log, the participants worked in groups to write a policy document on the same aspects of curriculum implementation dealt with in the individual logs. Each PLC was planned to allow all participants to discuss each aspect of implementation mentioned above in an online zoom session, then each wrote the individual log as an asynchronous task followed by the group log, which allowed for the participants to discuss their individual logs in their groups and write the group policy document together.

Validity of the data was reached by including a large database that covered all teacher- mentor participants. In fact, the study sample involved 81 participants and thus provided a sense of saturation to offer a level of coverage that made it possible to draw meaningful conclusions from the qualitative data (Dushnik and Tzabar Ben Yehoshua 2001). Additionally, confidence in the procedures was achieved by gathering data through sampling from three groups; thus, covering a wide range of research participants.

Reliability was addressed by applying rater-judgment to examine criteria and then analyze the log entries. They were analyzed and coded by the researchers for emerging themes using a grounded theory approach. Strauss and Corbin (1990) define four requirements for judging a good grounded theory: (a) it should fit the phenomenon deriving from varied data that is allied with the general field of knowledge; (b) it should provide understanding and clarity; (c) despite the data being comprehensive, it should provide a general view while including extensive variation that is abstract enough to be applicable to different contexts; finally, (d) it should state the conditions under which the theory applies, describing a reasonable basis for action.

Thus, in the current study, two rounds of analysis were conducted to fully understand the mediation sessions through the log entries. The first round was devoted to identifying recurring patterns and categories. This was done by each researcher independently and then discussed and compared while themes were identified and then categorized as they emerged using color-coding to highlight and manage log entries. Themes were then reviewed by an independent reader to determine whether they completely fit the data arising from the discourse in the implementation sessions. A comparison of the observations of both the researcher and the independent observer regarding the categories indicated a consistency rate of 90% and 92%, respectively.

It should be noted that the researchers took a central role as instructors of the PLCs. To eliminate possible sources of bias which might originate from the researchers' reasons for conducting the study, the motivating factors were viewed constructively and as a source of incentive leading to the desire to conduct research. This connects to Glaser and Strauss (1967) and Strauss and Corbin (1990: 42) who refer to the "theoretical sensitivity" of the researcher.

5 Findings

Seven themes were identified in the individual and group logs (see Table 1). Four themes related to RQ1 and three to RQ2. The first three focused on the perceptions and attitudes of the lead teachers, teacher-mentors and preservice teacher educators regarding their ability to mediate the implementation of the CEFR-based curriculum in the local context. The other three themes related to RQ2, which focused on how EFL lead teachers, teacher-mentors and preservice teacher educators construct their knowledge of the CEFR in a PLC.

Research Question One: Perceptions and attitudes

The first theme appears in the first set of individual logs and relates to the participants' confidence in understanding the curriculum document. It is worth initially considering the participants' feeling of confidence in understanding the curriculum and only then the research question which presents the second recurring theme, the reported ability of the participants to mediate the curriculum. After separately addressing these themes, a third theme emerged which demonstrated a connection between the participants' understanding of the document and their ability to mediate it. The fourth theme discusses the participants' reporting of the teachers' resistance to the new curriculum.

60 of the 81 (74%) participants in the three PLCs described their feeling as confident when talking about their understanding of the curriculum document.

> *"Generally, I think I have a good grasp of the rationale and principles of the curriculum."* (G1, T21)

> *"I believe that I have a good understanding of the 2020 Curriculum document."* (G2, T1)

> *"I feel that I understand the curriculum quite well. I use it when preparing lessons and assessment tasks together with teachers and I actively teach the curriculum in the teachers' college where I work."* (G3, T17)

In contrast, the remaining 27 described varying degrees of understanding and confidence. Their comments presented the spectrum and continuum of understanding with the addition of phrases which demonstrated some level of uncertainty.

> *"I feel I have some understanding of the new curriculum, but for me I feel I need more work to really understand and work with it.*" (G1, T6)

> *"I am not very familiar with the curriculum in depth."* (G2, T8)

> *"While I understand the general concepts, I haven't fully "internalized" them as of yet."* (G3, T4)

> *"I understand the curriculum well. However, there are many elements I don't know."* (G1, T20)

> *"I am familiar with and thoroughly understand* ***most*** *of its elements. So, what is the reason for my lack of confidence?"* (G1, T19)

> *"I am familiar with the new curriculum but definitely do not feel at the expert level yet."* (G2, T21)

It is then interesting to see if the same participants described a feeling of confidence to mediate the curriculum to others. 51 (63%) participants reported feeling completely confident about mediating the curriculum. This is eight less than those who reported they were confident with their understanding of the document. This seems to show that there is a correlation between being confident in their understanding of the curriculum and being able to mediate it to others.

> *"I feel well versed in the 2020 Curriculum and feel moderately confident in mediating it to others."* (G1, T11)

> *"Generally, I think I have a good grasp of the rationale and principles of the curriculum. I've now perfected a way of explaining the rationale and the relationship between the four skills and the four activities."* (G1, T10)

> *"Now that I feel more confident with the curriculum, it is easier to mediate it."* (G2, T2)

As highlighted by the previous comments of participants who feel confident about their understanding of the document and the following comments of participants who do not feel confident about

their understanding of the curriculum, all these participants connect their feeling of confidence in understanding the document to their ability to mediate it to others.

> *"I don't feel confident that I can explain all the information … I need to be more knowledgeable about everything. I have been trying to explain the curriculum with partial success."* (G1, T22)

> *"I am not very familiar with the curriculum in depth … I will be able to mediate it in a more precise way only after I understand it completely."* (G2, T8)

> *"I feel that I am not familiar enough with it in order to teach the pertinent information to my colleagues at school."* (G3, T4)

An additional theme that arose from the participants' perception of their ability to mediate the curriculum was the resistance of the teachers to whom they were mediating it. In the first course, 14 of the 30 participants, 47% mentioned the resistance of teachers to the introduction of the new curriculum. T18 said *"a teacher could not see the difference between the old and the new curriculum except the wording … The teachers don't understand. There was a lot of resistance. A feeling of continuous change."* (G1, T18) and T30 claimed; *"Teachers are sometimes resistant to change and automatically feel it will give them 'extra work'."* (G1, T30)

However, with the occurrence of each course the number of participants who related to the resistance of the teachers decreased. In the second course, 38% wrote about teacher resistance and in the third course, only 30% mentioned resistance in their individual logs.

Research Question Two: How teachers construct their knowledge

The next three themes relate to RQ2 dealing with how EFL lead teachers, teacher-mentors and preservice teacher educators construct their knowledge of the CEFR in a PLC. The individual logs at the beginning of the course were coded and compared with the final individual reflections written at the end of the course and common themes were found in the written discourse.

In the final reflection, participants reported an improved understanding of the document and an increased feeling of confidence to mediate the curriculum document to others. At the beginning of the courses, 51 out of 81 (63%) participants reported feeling confident to do so, but at the end of the course 67 out of the 81 (83%) participants stated that they felt more confident about the mediation. This is an increase of 20% of the percentage of participants reporting their feeling of confidence to mediate the curriculum. If more teachers reported a better understanding of the document and enhanced confidence to mediate it to other teachers, this is evidence that throughout the course they actively constructed knowledge.

> *"My current perception of my ability to mediate the curriculum to teachers has changed. I feel that I better understand the curriculum, its components, rationale and 'spirit', and I am able to convey the message to other teachers."* (G1, T19)

> *"I am amazed at my professional growth … I learned extensively about the mindset … and accumulated ideas for improving my teaching for my preservice students."* (G2, T10)

> *"I feel more confident than before this PLC. I feel that I should still read some parts over again to inculcate the message they convey, but all in all I feel I now have the necessary knowledge to mediate and advise my student teachers."* (G3, T16)

Analysis of the individual reflections yielded two dominant themes, which may explain the participants' increased self-efficacy regarding their understanding and ability to mediate the curriculum. The first relates to the framework of the course as a PLC, which encouraged different levels of interaction. The

following quotes relate to the benefits of working within the PLC framework; T19 said, *"The course was given in the form of a PLC with meaningful and interesting professional development. It was an excellent demonstration of a PLC which was run effectively and efficiently."* (G1, T19). T11 said, *"it also gave me a sense of a professional community … all along I felt like a 'lone soldier in the field'."* (G1, T11)

The second theme relates to the writing of four individual logs and four group logs. 67 out of 81 participants (83%) claimed that their enhanced ability to mediate the curriculum occurred due to the bottom-up process of writing the individual logs and then the social interaction in the group logs. This is the majority of the participants.

> *"I feel that this method of working with individual and then group logs was very empowering."* (G3, T24)

> *"The work on the group and individual logs was quite a unique experience for me. It was the first time that almost an entire course was designed around the logs which, in turn, forced us to constantly reflect and use meta-cognition on the work processes needed for mediating."* (G2, T5)

> *"Much of my learning came about by working with three experts in the field who each work in completely different settings."* (G3, T27)

> *"I found the Curriculum course last year to be quite theoretical and I gave my course not feeling very confident about the subject. […] I am happy to say that my initial impression of what this course would offer compared to how I felt at the end was vastly different. I initially thought that this would be another theoretical based course and how would I be able to apply this when mediating the "can-do" statements. I learnt so much from the other participants in the group and enjoyed the way the course was built – personal logs and then group logs. I found it very empowering."* (G3, T27)

These quotes emphasize the participants' feeling that while writing the individual and group logs, they underwent a process that allowed them to develop and improve their knowledge base individually and then as a group. The social interaction in their groups was particularly beneficial for knowledge construction.

83% of the participants felt that the course made a difference to their ability to mediate. The preservice teacher educators, however, referred to an additional benefit, the social interaction was very important for the transition of preservice teachers to the school system. It was an opportunity for these two often separate contexts of preservice and in-service instructors to connect and interact. *"The sessions highlighted the importance of joining forces in all layers of education* in *Israel to benefit all stakeholders."* (G2, T22)

6 Discussion

The individual and group reflective logs provided rich data with multiple recurring themes to answer the research questions. The perceptions and attitudes of the lead teachers, teacher-mentors and preservice teacher educators regarding their ability to mediate the implementation of the CEFR-based curriculum in the local context could be seen in the correlation found in the data between the participants' understanding of the curriculum document and their confidence to mediate it to others.

The additional theme of resistance is perceived by the participants throughout the courses, but the data shows a decrease in the participants reporting of the resistance. This, too, can be linked to the increase in the number of participants at the end of the course who reported an enhanced ability to mediate the curriculum (63% to 83%). This could indicate that when expert teachers feel more confident about their knowledge base, they are better able to mediate that knowledge to others, and thus, it results in a decrease in resistance from the mediatees, even if it requires a revised mindset.

Another reason for less resistance could be the introduction of materials and course books approved

by the Ministry of Education which were aligned with the mindset and concepts of the CEFR. The availability of example materials and model units could be a contributing factor to the continuum from policy to practice. However, further research is required to comprehend and explain the reason for the participants' decreased reports of resistance.

The second research question relating to the participants' changed perception of their ability to mediate the curriculum can be answered by addressing the remaining themes. A 20% increase of EFL lead teachers, teacher-mentors and preservice teacher educators who reported feeling confident or more confident to mediate the curriculum at the end of the PLC is evidence that knowledge construction of the CEFR in general and the curriculum documents in particular had occurred. Many participants (80%) recognize the importance of the bottom-up framework of the PLC which allowed for collaboration and sharing of ideas. Vygotsky's theory was helpful in explaining the process. Within the framework of the PLC, language was used as the tool that mediates the higher mental processes (Lapkin et al. 2010), and thus was referred to as a mediational tool expressed as verbalization or languaging, and could be described as "the process of making meaning and shaping knowledge and experience through language" (Swain 2006: 98). When the participants interacted and collaborated during the PLC, they were constructing knowledge through interaction and applying a collaborative dialogue (Swain and Lapkin 1998).

The additional theme highlighted by the participants' comments identifies the format of writing the individual and group logs as the contributing factor to the co-construction of their learning of the new curriculum. By first focusing on the individual perspective of a concept in the individual log and only then discussing the same concept in groups while writing a group log, the participants could initially clarify their personal beliefs and then construct a deeper and broader understanding informed by the perspectives of all group members. Furthermore, the unique integration of teacher-mentors facilitated a dialogue between two often-disconnected preservice and in-service contexts. Hence, the EFL lead teachers, teacher-mentors and preservice teacher educators collaborated and shared their difficulties and dilemmas.

7 Conclusion

This study explored a model of implementing a localized CEFR curriculum in two stages, inspired by a social-cultural perspective in planning, design and activation. The curriculum implementation process can be seen in Figure 1. The ripple effect was clearly portrayed in the different stages of implementation. In the middle of the vortex, curriculum document policy was adapted from the CEFR. In the first year, the document was introduced to the field with PD courses as seen by the light blue arrows. The participants in these courses grappled with the terminology and the mindset while changing their thinking about practice, as they compared the old to the new. The second stage of implementation required a different approach to PD. To this end, it was found useful to talk about curriculum implementation applying a PLC format (dark arrows) from Vygotsky's theoretical perspective. Moreover, it was found effective for grappling the meaning of the curriculum with a group of English lead-teachers and teacher-mentors whose role was to mediate their learning to other English teachers by adopting the concept of mediation, especially to EFL language educators. The PLC framework was adopted because it allowed the participants to collaborate and interact to develop their ability to mediate the curriculum to themselves and to others. The aim is clearly to reach the practice level whereby the curriculum is implemented by teachers in the field. The ripple effect metaphor is the most apt, considering the desire to disseminate the curriculum to a broader audience and to ensure that it becomes more than a policy document.

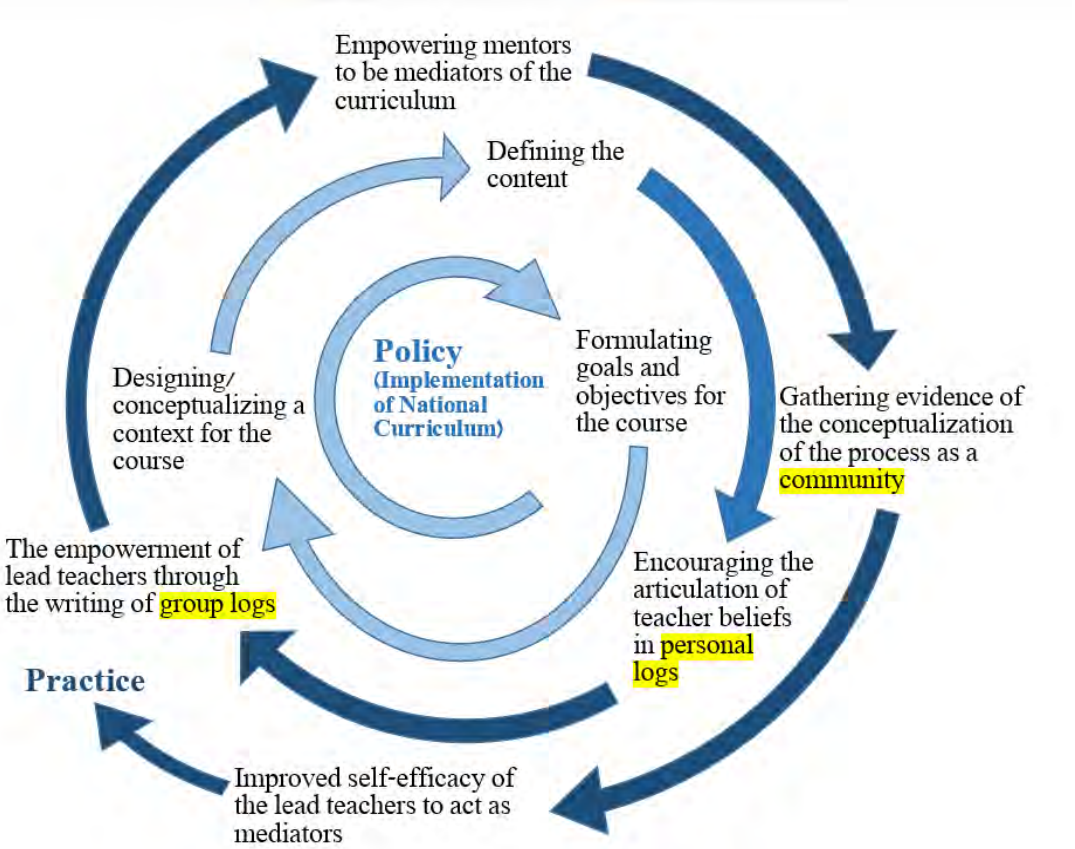


Figure 1. Curriculum implementation process

8 Further research

Based on the results of this study, a third stage of implementation is required in addition to further research. From this study, we cannot equivocally claim that the principles of the CEFR have influenced actual teaching practices. Based on the individual logs, the mindset of the expert teachers and the teachers they work with has, indeed, shifted. However, this is not evidence of a direct impact on teachers' planning and students' learning in the classroom. Lead teachers, teacher-mentors and preservice teacher educators now need to facilitate courses with teachers in the field to prepare materials according to the guidelines of the CEFR-based curriculum. Such a process will allow practical discussions about classroom teaching and will empower these experts to guide the teachers to adapt current materials in addition to preparing materials suited to their individual classroom context. The case study protocol could be a suitable framework for Stage 3 implementation. There is a need to observe the classroom setting in order to evaluate the ripple effect and that policy is disseminated to practice.

9 References

Ambrosetti, Angelina. 2014. Are you ready to be a mentor?: Preparing teachers for mentoring pre-service teachers. *Australian Journal of Teacher Education (Online)*, *39*(6), 30-42.

Becher, Ayelet & Lily Orland-Barak. 2016. Integrating social activity theory and critical discourse analysis: A multilayered methodological model for examining knowledge mediation in mentoring. *International Journal of Qualitative Studies in Education*, 29(4), 498-519.

Borg, Simon. 2006. The distinctive characteristics of foreign language teachers. *Language teaching research*,10(1), 3-31.

Corbin, Juliet M. & Anselm Strauss. 1990. Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative sociology*, *13*(1), 3-21.

Council of Europe. 2020. Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume, *Council of Europe Publishing, Strasbourg*. https://www.coe.int/lang-cefr (accessed 22 October 2022).

Donato, Richard. 1994. Collective scaffolding in a second language. *Vygotskian approaches to second language research*, 33-56.

DuFour, Richard. 2004. What is a" professional learning community"? *Educational leadership*, 61(8), 6-11.

Dushnik, Liron., & Tsabar Ben Yehoshua, Naama. 2001. Ethics of Qualitative Research, In Tsabar Ben Yehoshua (Ed.) *Traditions and Currents in Qualitative Research*. Israel: Dvir Press, 343-367. (in Hebrew)

Feiman-Nemser, Sharon. 2001. *From preparation to practice: Designing a continuum to strengthen and sustain teaching.*

Frawley, William, and James P. Lantolf. 1985. Second language discourse: A Vygotskyan perspective. *Applied linguistics*, 6(1), 19-44.

Fullan, M. (Ed.). 2003. *The moral imperative of school leadership*. Corwin Press.

Israeli Ministry of Education. 2020. Professional Framework for English Teachers 2020. https://meyda.education.gov.il/files/Mazkirut_Pedagogit/English/framework 2020.pdf (accessed 22 October 2022).

Gánem Gutiérrez, Adela. 2008. Microgenesis, method and object: A study of collaborative activity in a Spanish as a foreign language classroom. *Applied linguistics*, 29(1), 120-148.

Gánem Gutiérrez, Adela. 2013. Sociocultural theory and second language development. *Contemporary approaches to second language acquisition*, *9*, 129.

Genç, Zübeyde Sinem. 2016. More practice for pre-service teachers and more theory for in-service teachers of English language. *Procedia-Social and Behavioral Sciences*, 232, 677-683.

Glaser, B., & A. Strauss. 1967. The discovery of grounded theory Aldine Publishing Company. *Hawthorne, New York*.

Kim, Min-Young & Ian A.G. Wilkinson. 2019. What is dialogic teaching? Constructing, deconstructing, and reconstructing a pedagogy of classroom talk. *Learning, Culture and Social Interaction*, 21, 70-86.

Lantolf, James P. 2006. Sociocultural theory and L2: State of the art. *Studies in second language acquisition*, 28(1), 67-109.

Lantolf, James P. & Steven L. Thorne. 2006. *Sociocultural theory and genesis of second language development*. Oxford: Oxford University Press, 2006.

Lantolf, James P & Gabriela Appel (eds.). 1994. *Vygotskian approaches to second language research.* Greenwood Publishing Group.

Lefstein, Adam, Nicole Louie, Aliza Segal & Ayelet Becher. 2020. Taking stock of research on teacher collaborative discourse: Theory and method in a nascent field. *Teaching and teacher education*, 88, 102954.

Ohta, Amy Snyder. 1995. Applying sociocultural theory to an analysis of learner discourse: Learner-learner collaborative interaction in the zone of proximal development. *Issues in applied linguistics*, 6(2), 93-121.

Pang, May. 2019. Developing core practices for EFL/ESL teaching: A framework for methodology course design. *Tesol Quarterly*, 53(1), 258-273.

Popp, Jacquelynn S., and Susan R. Goldman. 2016. Knowledge building in teacher professional learning communities: Focus of meeting matters. *Teaching and Teacher Education,* 59, 347-359.

Scardamalia, Marlene, and Carl Bereiter. 2003. Knowledge building environments: Extending the limits of the possible in education and knowledge work. *Encyclopedia of distributed learning,* 269-272.

Shohamy, Elana. 2001. Democratic assessment as an alternative. *Language testing*, 18(4), 373-391.

Stoll, Louise, Ray Bolam, Agnes McMahon, Mike Wallace & Sally Thomas. 2006. Professional learning communities: A review of the literature. *Journal of educational change,* 7(4), 221-258.

Swain, Merrill. 2006. Languaging, agency and collaboration in advanced second language proficiency. *Advanced language learning: The contribution of Halliday and Vygotsky*, 95-108.

Swain, Merrill & Sharon Lapkin. 1998. Interaction and second language learning: Two adolescent French immersion students working together. *The modern language journal*, 82(3), 320-337.

Swain, Merrill & Sharon Lapkin. 2000. Task-based second language learning: The uses of the first language. *Language teaching research*, 4(3), 251-274.

Swain, Merrill & Yuko Watanabe. 2013. Languaging: Collaborative dialogue as a source of second language learning. *The encyclopedia of applied linguistics*, 3218-3225.

Tillema, Harm & Lily Orland-Barak. 2006. Constructing knowledge in professional conversations: The role of beliefs on knowledge and knowing. *Learning and instruction*, 16(6), 592-608.

van Schaik, Patrick, Monique Volman, Wilfried Admiraal & Wouter Schenke. 2019. Approaches to co-construction of knowledge in teacher learning groups. *Teaching and teacher Education*, 84, 30-43.

Vygotsky, Lev Semenovich & Michael Cole. 1978. *Mind in society: Development of higher psychological processes*. Harvard University Press.

Watanabe, Yuko & Merrill Swain. 2007. Effects of proficiency differences and patterns of pair interaction on second language learning: Collaborative dialogue between adult ESL learners. *Language teaching research*, 11(2), 121-142.

Williams, Lawrence, Lee B. Abraham & Eduardo Negueruela-Azarola. 2013. Using concept-based instruction in the L2 classroom: Perspectives from current and future language teachers. *Language Teaching Research*, 17(3), 363-381.

10 Biographies

Dr Tziona Levi is the director of the department of languages and the chief inspector for English Language Education at the Ministry of Education in Israel. Her fields of interest include applying dynamic assessment (DA) to various EFL teaching and learning contexts and promoting plurilingual/pluricultural agendas in school. She has established 68 teacher Learning Communities nationwide to promote professional development among English language teachers and English oral proficiency in schools. She headed the process of rewriting the Israeli national English curriculum to be aligned with international standards, mainly the CEFR.

Simone Duval is the District English Superintendent for Jerusalem, Israel. She has 29 years of experience as an EFL school teacher in Israel. She is a teacher trainer in Education courses and has developed online courses for teachers and students. In addition, she coordinated the implementation of the revised Israeli English curriculum.

Appendix A

Individual Logs

LOG 1 Curriculum Implementation

Name:	Date:
Questions	My log entry
Write up to 100 words. 1. How well do I understand the curriculum? 2. What actions should I take to enhance my understanding of the curriculum and my ability to mediate it to others? 3. What are my strengths in this area?	
Write up to 100 words. 1. Consider all of the interactions you have had with teachers or student-teachers regarding implementation of the new curriculum. Describe two of the interactions. 2. Give one example of how you mediated the mindset of the new curriculum during these interactions. 3. If you have not yet interacted with teachers or student-teachers, suggest how this should be done within your professional context.	

LOG 2 Backward Design

Name:	Date:
Questions	My log entry
Write 50-100 words. 1. How well do I understand the concept and application of backward design? 2. What actions should I take to empower teachers to implement backward design in their planning?	
Write up to 200 words. 1. Ask the teachers in your course to prepare the backward design template. Explain with examples. If you are not teaching a course, give examples from your experience working with teachers in the field. 2. What were the issues you encountered with the teachers and how were they solved? 3. What were the difficulties that arose during the teachers' process? 4. What would you do differently next time to ease the process?	

LOG 3 Ministry Resources

Name:	Date:
Questions	My log entry
Write 100-150 words. 1. What do you know about the Ministry of Education's recorded lessons and teaching units and the English Inspectorate's emphasis on them as a teaching resource other than the online teaching during the COVID 19 era? 2. To what extent do you see these resources as a tool for the continued implementation of the curriculum?	
Write up to 200 words. 1. Explain how you would or did incorporate these resources into your courses or counseling. 2. Give specific examples of how you think this rich and extensive resource can be used to enhance the noticing of the teachers with regards to the implementation of the Curriculum.	

LOG 4 Simulations

Name:	Date:
Questions	My log entry
Write 100-150 words. 1. Describe an experience you have had in a simulation center or session as a participant or observer. 2. How did you feel and what was your take away? Explain with specific examples.	
Write up to 200 words. 1. As a counselor or PLC leader what considerations did you take into account when you prepared the session? (If you have never taken a group, discuss considerations you feel need to be addressed.) 2. Possible considerations; logistics, scenarios, expected outcomes. 3. Describe the feedback session you had with your group after the simulation (If you have not taken a group then answer this question instead of the one above; Which questions would you ask the group in a post simulation session.) 4. What lessons were learned by the participants and you as the leader? (Or define the role of the PLC leader or course instructor that takes a group to a simulation center or simulation session.)	

Appendix B

Group Logs

LOG 1A

In groups of 5 share your individual LOGS with the members of your group.

You are now the member of an advisory committee to the English Inspectorate.

They have asked your group to suggest a framework for the effective implementation of the new curriculum.

Committee members: _____

What is the issue? (Describe the need for an effective framework for the implementation of the revised curriculum in 5 sentences).

This framework takes into consideration:

1.

2.

3.

Describe your suggested framework for the implementation of the revised English curriculum in Israel. Outline the steps and relate to the various stakeholders. (Write 200-250 words.)

LOG 2A

In your groups, share personal LOG 2 with the members of your group.

In your advisory committee, discuss the following.

Suggest a framework for the effective implementation of backward design as a tool to plan towards achieving learning outcomes (can do statements).

Committee members: _____

1. Explain why backward design planning is an important framework for teachers to understand and adopt within the new curriculum and especially now during this time of blended teaching and learning.
2. Suggest the stages (smaller units) of how you would present the backward design framework to the teachers. (Up to 4 stages)
3. These suggestions take into consideration: (For example; synchronous, asynchronous counseling, teacher resistance, preservice/ in-service training etc.)

LOG 3A

In your groups, share personal LOG 3 with the members of your group.

In your advisory committee, discuss the following.

Suggest a framework for the practical use of the filmed lessons as a tool to improve the teachers' understanding of the curriculum.

Committee members: _____

Explain how you would present the filmed lessons to the teachers. Which guiding or high leverage questions connected to the language activities would you ask the teachers?

- Suggest the stages (smaller units) of how you would present the filmed lessons to the teachers. (Up to 4 stages)
- These suggestions take into consideration: (For example; preparing a viewing tool)

LOG 4A

In your groups, share personal LOG 4 with the members of your group.

In your advisory committee, discuss the following.

Suggest a framework for the practical use of the simulations as a tool to improve the teachers' understanding of the curriculum and more effective teaching practices.

Committee members: _____

- Define the aims and desired outcomes of taking a group to a simulation center.
- Considering the enormous expense of simulations, your committee has been asked to suggest an alternative framework that will achieve the same or similar aims.
- Even in the post Corona era, distance learning will be a viable alternative in many situations. Compare and contrast the online simulation to the conventional face to face simulations at the center.
- Write a possible scenario that can be useful for a simulation with your teachers on implementing the curriculum.

Guidelines for writing a scenario:

4. Who are the people involved in the scenario? (teachers, coordinator, principal, parent, ...)
5. What is the topic of the scenario? (Implementation of backward design/ resistance, ...)
6. Describe the scenario and include the point of view of the individuals in the scenario and a difficulty or conflict.
7. Consider the desired outcome of the scenario.

News

We are happy to receive reports on events, conferences or projects concerning CEFR, CEFR/CV or Language Portfolio related themes.

COE https://www.coe.int/en/web/common-european-framework-reference-languages/home

ALTE Association of Language Testers in Europe https://www.alte.org/, https://www.alte.org/Events

ALTE ALTE Language Assessment Fit for the Future. Madrid, April 26-28, 2023

ALTE CEFR SIG https://www.alte.org/CEFR-SIG

EALTA European Association for Language Testing and Assessment https://www.ealta.eu.org/

EALTA Annual conferences https://www.ealta.eu.org/conferences.htm

19th EALTA Conference: When: **June 13-18, 2023**, Where: **Helsinki, Finland**

EALTA CEFR SIG https://www.ealta.eu.org/sig.html

Pre-conference workshops: June 13-15, 2023

ECML European Center of Modern Languages

VITbox project: CEFR Companion Volume implementation toolbox

https://www.ecml.at/ECML-Programme/Programme2020-2023/CEFRCompanionVolumeimplementationtoolbox/tabid/4299/language/en-GB/Default.aspx

ADiBE Project https://adibeproject.com/

Link to the Project Card: https://ec.europa.eu/programmes/erasmus-plus/projects/eplus-project-details/#project/2018-1-ES01-KA201-050356

The ADiBE project brings together key figures with ample experience in the field of Content and Language Integrated Learning (CLIL) from six countries (Spain, Austria, Finland, Germany, Italy, and the UK).

JALT CEFR & LP SIG

- CERF & LP SIG: https://cefrjapan.net
- CEFR Journal: https://cefrjapan.net/journal
- CEFR events: https://cefrjapan.net/events

Book-Reviews

If you are interested in reviewing CEFR-related publications, books, and textbooks, we have several offers. This is for academic and educational purposes not for commercial interests. Please contact the editors.

CEFR JOURNAL—RESEARCH AND PRACTICE
VOLUME 5

Title: **CEFR Journal—Research and Practice**

Type: Online Journal
URL: https://cefrjapan.net/publications/journal
Contact: journal@cefrjapan.net

Edited by: Japan Association for Language Teaching (JALT)
CEFR & Language Portfolio SIG
Fergus O'Dwyer (editor)
Maria Gabriela Schmidt (coordinator, editor)
Morten Hunke (liaison officer, editor)
Alexander Imig (treasurer, website editor)

ISSN: 2434-849X
DOI https://doi.org/10.37546/JALTSIG.CEFR

Publication Statement of the CEFR Journal

All work published in the CEFR journal is original work, with contributions from the named authors.

The work published in the CEFR journal is open access and licensed under a Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) license.

Authors of articles are allowed to retain publishing rights and hold the copyright without restrictions. The CEFR Journal actively checks for plagiarism using Turnitin. The CEFR Journal does not charge article processing charges or submission charges.

Any issues regarding complaints and appeals, conflicts of interest / competing interests; ethical oversight; post-publication discussions and corrections should be directed to journal@cefrjapan.net

The administration costs of the CEFR Journal are covered by the JALT CEFR & Language Portfolio SIG; the CEFR Journal has no paid advertisements or other revenue sources.

In the event that the CEFR Journal is no longer published, electronic backup and preservation of access to the journal content shall be made available via CLOCKSS.

Submissions (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.

Currently, we have a new Call for abstracts out. Due to current necessities and demand, we are looking to give your experiences with online, remote, and e-learning in conjunction with the CEFR, the CEFR/CV, or portfolio work the spotlight it deserves. In these months many practitioners are accruing valuable best and potentially also worst practice experience. We would like to offer a forum to share such valuable insights in future volumes. Until 28 February 2023 we are looking for abstracts at: journal@cefrjapan.net.

Please contact the editors with any queries and submit to:

journal@cefrjapan.net

Guidelines

Submission:	28 February 2023.
Contributions:	Articles (research), reports (best practice), news (work in progress), research notes, book reviews
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

Reviewer guidelines:

We ask all reviewers to make every reasonable effort to adhere to the following ethical guidelines for CEFR Journal—Research and Practice articles they have agreed to review:

1. Reviewers must give unbiased consideration to each manuscript submitted for consideration for publication, and should judge each on its merits. Since, we employ a double-blind review, the text you have been provided with ought to have no reference to race, religion, nationality, sex, gender, seniority, or institutional affiliation of the author(s). Please, notify us immediately where any such information is still detectable in the anonymised text you received.
2. 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.
3. Reviewers are strongly advised to keep the peer review process confidential; information or correspondence about a manuscript should not be shared with anyone outside the peer review process.
4. Reviewers should provide a constructive, comprehensive, evidenced, and appropriately substantial peer review report. For your convenience, we are providing you with a 'reviewing matrix' you may choose to use at your own discretion. We would also like to kindly ask you to provide us in the journal editorial team with a final overall assessment of the text's publication potential—please, see

bottom of this document.

5. Reviewers must avoid making statements in their report, which might be construed as impugning any person's reputation.
6. Reviewers should make all reasonable effort to submit their report and recommendation in a timely manner, informing the editor if this is not possible.
7. Reviewers should call to the journal editor's attention any significant similarity between the manuscript under consideration and any published paper or submitted manuscripts of which they are aware.

Author instructions:

Adapted version of deGruyter Mouton guidelines for Language Learning in Higher Education (CercleS) and style sheet.

CEFR Journal—Research and Practice

Japan Association for Language Teaching (JALT)
CEFR & Language Portfolio SIG (CEFR & LP SIG)

ISSN 2434-849X