DiSlaw – Didaktik slawischer Sprachen
ISSN: 2960-4117
dislaw.at
2023, Special Issue / Sonderausgabe
DOI: 10.48789/2023_special_issue3

Benjamin Kremmel, Universität Innsbruck, Austria

Kathrin Eberharter, Universität Innsbruck, Austria

Eva Konrad, Universität Innsbruck, Austria

Elisa Guggenbichler, Universität Innsbruck, Austria

Doris Moser-Frötscher, Universität Innsbruck, Austria

Viktoria Ebner, Universität Innsbruck, Austria

Carol Spöttl, Universität Innsbruck, Austria

The CEFR Companion Volume: Opportunities and challenges for language assessment

Over the last two decades, the Common European Framework of Reference for languages (CEFR) has become the most influential tool of language policy-making in Europe and beyond. The publication of the companion volume (CEFR-CV) constitutes a new milestone for teaching, learning and assessing languages, and is a most timely reaction to common criticism of the framework. In addition to new scales, descriptors and competence levels, the CEFR-CV introduces new modalities and broadens the scope for mediation and plurilingual/cultural communication, thereby updating and extending previous construct definitions for increasingly digitized and diverse societies. Despite the CEFR's major impact on the language testing industry, there is thus far scarce literature on how to operationalize the CEFR-CV for assessment with the expanded framework. In addition to the huge potential for innovative assessment tasks and formats, this raises questions with regard to construct definitions, task development, test quality assurance, and rating practices. This paper will focus on six noteworthy innovations of the CEFR-CV and discuss the opportunities and challenges for assessment: (1) departure from the native-speaker norm, (2) stronger consideration of digital communication, (3) interlingual mediation, (4) intralingual mediation, (5) phonological awareness, and (6) the provision of richer descriptions of lower-level learner competencies.

Keywords: CEFR, companion volume, CEFR-CV, assessment, language testing

1. Introduction

Over the last two decades, the Common European Framework of Reference for Languages (CEFR) has become the most influential tool of language policy-making in Europe and beyond (e.g., Figueras, 2012). As an "indispensable reference point for all aspects of second and foreign language education" (Little & Figueras, 2022, 15), it has had particular impact in the area of language assessment, with many placement, proficiency certification, educational monitoring, and school-leaving examinations across the globe referring to levels and descriptors from the framework for their exam design and score interpretation. The recent publication of the companion volume CEFR-CV to the CEFR (CEFR-CV; Council of Europe, 2020) therefore constitutes another milestone for teaching, learning and assessing languages. The CEFR-CV is a reaction to common criticism of the framework (cf. Bärenfänger et al., 2017; Deygers, 2021 for an overview) and represents another important step in advancing the communicative, action-oriented approach in language education. As such, it is undoubtedly bound to strongly affect language assessment as well, as it updates and extends previous construct definitions for increasingly digitized and diverse societies (Bärenfänger et al., 2017; Deygers, 2021). Despite the likely wide-ranging impact and the CEFR-CV project being more than 10 years in the making by now, there is sparse literature to date on how to operationalize the CEFR-CV for assessment with the expanded framework (British Council et al., 2022). However, the new scales and descriptors indeed "may fuel innovation in language education" (North, 2022, 51) as they provide huge potential for both developing novel assessment tasks and formats but also for discussing new questions with regard to construct definitions, test quality assurance, and rating practices (e.g., Kantarcıoğlu, 2022; Lenz, 2022). This conceptual paper will focus on six of the most noteworthy innovations of the CEFR-CV, and will discuss the opportunities and challenges these updates of the original CEFR may provide for assessment purposes: (a) departure from the native-speaker norm, (b) revisions to the phonological control scale, (c) stronger consideration of digital communication, (d) interlingual mediation, (e) intralingual mediation, and (f) the provision of richer descriptions of lower-level learner competencies.

2. Departure from the native-speaker norm

For decades, the notion of the native speaker has been an influential concept in language learning and assessment. Native-like proficiency has been considered the goal of language learning (Little, 2022). However, the notion of the native speaker has been controversial, to say the least, even when the CEFR was first published in 2001 (Council of Europe, 2020). There are two main arguments against the dominance of the native speaker notion. First, as a goal for language learning, it is unreasonable and mostly unattainable (McNamara & Shohamy, 2016). Second, it presupposes that there is one generally accepted standard used by L1 speakers, whereas in fact, most L1 users do not speak the standard variety themselves. In contrast to actual native speakers with their varied linguistic profiles, the conceptual native speaker, who defines the standard, is an idealization (Davies, 2012). Refuting the notion of the native speaker, however, does not mean that there should be no standards in language learning. As Widdowson (1994) acknowledged, referring to international communities using English as a shared language, "it is helpful, to say the least, to have a standard language at your disposal. But you do not need native speakers to tell you what it is" (Widdowson, 1994, 382).

The CEFR (Council of Europe, 2001) has been understood by a wide audience to uphold the notion of the native speaker. There is some indication that this may not have been its intention: the 2001 version contained references to interculturalism and plurilingualism (Bärenfänger et al., 2017; North, 2021), and explicitly stated that the level C2, the highest level described, "is not intended to imply native speaker or near native speaker competence" (Council of Europe, 2001, 36). In fact, North calls this one of the "myths that have developed around the CEFR" (North, 2022, 53). However, precisely in using this concept, even if just for demarcation purposes, the CEFR authors might have contributed to the emergence of this 'myth', which can most likely be traced back to the wording of many level descriptors and has been taken up beyond the educational sector, e.g., in job advertisements which frequently require 'native' or 'near-native' proficiency. The scales, undoubtedly the most influential part of the CEFR (Alderson, 2007; Deygers et al., 2018), referred to the native speaker explicitly in individual descriptors as well as the titles of entire scales like Understanding conversation between native speakers or Understanding a native speaker interlocutor. Such titles seemed to imply that L2 speakers only communicate with native speakers (McNamara & Shohamy, 2016). In the DIALANG self-assessment scales (Appendix C in the CEFR 2001, Council of Europe, 2001) which are no longer included in the CEFR-CV (2020, 232), several descriptors referred to the native speaker: e.g.

the C2 writing descriptor "I can write so well that native speakers need not check my texts" or the C1 listening descriptor "I can keep up with an animated conversation between native speakers" (ibid., 234). Whether intended or not, using this terminology meant that the CEFR was understood even by well-informed members of the research community to uphold the "ideology of the privilege of the native speaker" (McNamara, 2014, 229).

In the 2020 CEFR-CV, consequently, 16 descriptors at B and C levels have been revised to remove the term native speaker. Another maybe less noticed but nevertheless important change is visible in the scale Understanding an interlocutor, where the native speaker used to feature in the scale title and in the C2 descriptor, the term "non-standard accent or dialect" was substituted with "less familiar variety" (Council of Europe, 2020, 73). In the same scale, the B2 descriptor "Can understand in detail what is said to him/her in the standard language" was extended to include "[…] or a familiar variety" (ibid., 73). Further, the original scale for phonological control, which reinforced the wide-spread belief that accent was a sign of poor phonological control (Council of Europe, 2020), was completely rewritten to focus on "intelligibility rather than any native speaker model and admit that even obviously C2 speakers frequently retain an accent" (North, 2021, 12).

Removing the word native speaker and explicitly referring to "familiar" (B1/B2) and "less familiar" (C2) varieties has implications for assessment. "Non-standard" implies a generally accepted standard, whereas "less familiar" signals space for localization of CEFR-related assessment. What is more or less familiar will very much depend on the context of the assessment in question, and this formulation invites test developers to critically reflect upon the varieties of language their candidature may encounter. Of course, this needs to be done carefully and with test use in mind. Any test intended for international use will need to be more universal in approach than a test for a specific context.

The explicit departure from the native speaker norm is most likely to have consequences for the assessment of speaking and listening. Concerning listening, the CEFR-CV may offer a chance for test developers to widen the scope of speakers used in sound files for the assessment of B- and C-level listening. This may include not only featuring more or less familiar varieties, but could also extend to speakers with other L2 accents (Harding, 2011). In today's globalized, multilingual world, communication between language learners in the target language is at least as likely and frequent as communication between language learners and L1 speakers. Excluding L2 accents completely may therefore have negative consequences for the validity of any listening assessment. In speaking, test developers need to critically evaluate their rating scales against the new scale for

phonological control, with its focus on sound articulation and prosodic features. Also, rater training needs to raise awareness about raters' reactions to candidates' accents as will be discussed in the following section.

3. Phonological control

Of those scales initially published in the first iteration of the CEFR (Council of Europe, 2001), the scale describing phonological control is the one that saw the most drastic revisions and additions in the CEFR-CV (Council of Europe, 2020), while most of the scales describing linguistic competences were only revised minimally (e.g., general linguistic range) or not at all (orthographic control). As the comment in the CEFR-CV explains, the phonological control scale was the "least successful" (Council of Europe, 2020, 133) scale when calibrated in the empirical foundation work. This suggests that participants in the calibration project had difficulties with applying the draft descriptors, and indicates fundamental theoretical issues in the operationalization of the construct.

The original phonological control scale consisted of five descriptors (A1–C1/C2), one per level, which combined operationalizations of sound level phonetics and paralinguistic phenomena (stress, rhythm, and intonation) (Council of Europe, 2001, 117). As the CEFR came to be used more widely, a handful of issues emerged which highlighted fundamental weaknesses of this scale (Piccardo, 2016). First, it was critiqued for its general lack of detail and "skeletal descriptors" (Galaczi et al., 2011, 67). While scales are understood to be somewhat vague by necessity, raters struggled to match the descriptors with actual speaking performances (e.g., the attributes "natural" [B2] and "noticeable" [A2] in Harding's [2016] investigation). Another shortcoming was a perceived inconsistency in that some features (i.e., intonation) only appeared at certain levels of the scale (Harding, 2016). Focusing on salient features as they present themselves on particular levels had helped to keep the descriptors brief (North, 2014, as cited in Harding, 2016, 21–22), but only mentioning features sporadically posed a problem to rating scale developers and teachers who are interested in how certain abilities progress in language learners. Finally, attributes such as "foreign" and "natural" continued a teaching tradition that sees "accent as [...] a marker of poor phonological control" (Council of Europe, 2020, 133). This view is unhelpful in language pedagogy as it disregards context and individual learner needs and is now considered out of step with concepts such as intelligibility and accentedness, which have replaced previous constructs in speaker perception research (cf. section above).

As can be seen from this overview, the original phonological control scale fell short in capturing the underlying constructs in a way that satisfied practitioner and researcher needs (e.g., Harding, 2017). It is therefore not surprising that the CEFR-CV published three completely new scales instead of just revising the original scale: overall phonological control (which consists of entirely new descriptors and replaces the original scale), sound articulation and prosodic features (cf. Table 1).

The phonological control scale now provides a more systematic and detailed operationalization of four key concepts (cf. Table 1). The B2 level now also appears more realistically attainable than in the original scale ("accent tends to be influenced by the other language(s) they speak, but has little or no effect on intelligibility", [Council of Europe, 2020, 134]). The other two scales provide additional detail, while also clarifying the boundaries between the different components of the pronunciation construct. With this added level of detail, the CEFR-CV has the potential to serve as a more solid basis for developing and revising rating scales in many assessment contexts. Nevertheless, one shortcoming that rating scale developers will have to continue to overcome is the fact that fluency and pronunciation are still kept apart, even though raters often struggle to differentiate the two (Harding, 2016).

Phonological control	
Scale title	Key components
Overall phonological control	• intelligibility • influence from other languages spoken • control of sounds • control of prosodic features
Sound articulation	• clarity and precision of pronouncing sounds in the language
Prosodic features	• control of stress, intonation, and rhythm • ability to exploit and vary stress and intonation to highlight their message

Table 1: Overview of scales operationalizing phonological control
(based on Council of Europe, 2020, 133)

The revisions in the current CEFR-CV with regards to the area of pronunciation can also be regarded as evidence for the success of a prolific branch in current applied linguistics. There has indeed been a rise of interest in L2 pronunciation s (Isaacs & Harding, 2017; Munro & Derwing, 2011). While the teaching of

pronunciation has been mostly marginalized and teachers often left to their own devices, it is likely that this new scale in the CEFR-CV will have a noticeable impact on foreign language pedagogy. Regarding assessment, the novel scale necessitates revisions of existing rating scales for speaking, in which pronunciation often features prominently, and possibly a recalibration of examiner expectations and standards.

4. Communication in the digital age

With the advance of digital technology and particularly over the last years due to the Covid-19 pandemic, communication and interaction has increasingly moved to online mode for the public, educational, occupational, and also personal domains. Examinations have already – for obvious reasons of authenticity – replaced letters with e-mails, and included tasks requiring test takers to write blog posts or blog comments (the Austrian school-leaving exam is a case in point, cf. Bundesministerium für Bildung, Wissenschaft und Forschung [BMBWF], w.d.). Aside from featuring updated text and task types, entire examinations and assessment systems have moved or have had to move to online mode, largely necessitated by the pandemic (cf. e.g., Isbell & Kremmel, 2020).

Although the 2001 version of the CEFR was criticised for its lack of representation of online communication and interaction (North, 2021), it has to be stressed that it was indeed already represented to some extent. For instance, e-mail and online computer conferences were given as examples for correspondence and interaction, respectively (Council of Europe, 2020, 82). Further trends of the digital age were in fact foreseen by the 2001 version, for example the advent of interactive human-machine communication in the public, occupational, educational, and personal domains (Council of Europe, 2001, 82), and the convergence of electronic interaction to "'real time' interaction" (ibid., 92). The authors also stressed the need for learner exposure to authentic L2 language use through computer programmes and online communication (e.g., "international computer networking of schools, classes and individual students" [Council of Europe, 2001, 145]). However, and understandably so given the timeframe for the project, the descriptors themselves did not explicitly refer to computer-mediated or online communication. In the CEFR-CV, however, two new descriptor scales for online interaction were added:

1. Online conversation and discussion, and
2. Goal-oriented online transactions and collaboration

According to the CEFR-CEFR-CV, "[b]oth these scales concern the multimodal activity typical of web use, including just checking or exchanging responses, spoken interaction and longer production in live link-ups, using chat (written spoken language), longer blogging or written contributions to discussion, and embedding other media" (Council of Europe, 2020, 25). The CEFR-CV states that successful online communication necessitates more redundancy, checking whether the message was understood correctly, reformulation to facilitate understanding and resolve misunderstandings, and handling emotional responses (Council of Europe, 2020, 84). There were also additions within descriptors for other language activities to include reference to language use in the online environment, for instance discussion forums, blogs, postings, confirmation of a booking or online purchase, messages sent via social media as examples for visual reception (*Reading correspondence*, ibid., 54–55), websites (*Reading for orientation*, ibid., 55–56), instructions for installing new technology (*Reading instructions*, ibid., 58), and chat and forum (*Identifying cues and inferring*, ibid., 60). These new scales for online communication have yet to be taken up on a larger scale by test developers. Cinganotto (2019) reports on a pilot project in Italy which showed the usefulness of the *Online interaction scales* as a basis for designing collaborative online communication tasks in an upper-secondary school context. The tasks were met with positive feedback by students and teachers alike.

However, these additions have also not gone without criticism. In their more critical discussion of the online communication scales, Bärenfänger et al. (2017) see an undue focus solely on the medium of online communication and criticise a lack of task- and situation-oriented descriptions of language ability. They argue that the requirements for production and reception of language in the online setting, such as features of spoken language in written texts, or the use of visual elements like emoticons remain unconsidered. However, upon closer inspection, some of these points can be easily rebuked. The CEFR-CV does point out the increasingly multimodal nature of online interaction by stating that "[a] rigid separation between written and oral does not really apply to online transactions" (Council of Europe, 2020, 86). It also exemplifies the incorporation of "symbols, images and other codes […] to convey tone, stress, and prosody" (Council of Europe, 2020, 84). In the descriptors for online interaction, the CEFR-CV offers detailed descriptions of what language learners can do when communicating in different online settings but also, contrary to Bärenfänger et al.'s (2017) critique, specific situations (e.g., "live, online professional or academic discussion", *Online*

conversation and discussion, Council of Europe, 2020, 85) and roles (e.g., "service role", *Goal-oriented online transactions and collaboration*; Council of Europe, 2020, 86).

For language testing, the CEFR-CV's new and amended descriptors highlight the vast array of online communicative situations and map out the language progression for online interaction, thus furnishing a frame of reference for better developing level-specific test tasks for online communication and interaction. With ongoing technological developments, language tests will more and more authentically be able to tap into these situations, for example through the use of chatbots for goal-oriented transactions. The CEFR-CV enables and encourages test developers to explore a new range of authentic task types, will likely further promote the trend towards digital exam delivery, and will allow rating scale developers and users to better reflect target language use in the digital age, which increasingly includes both cross-linguistic and intralinguistic mediation, as will be discussed in the following two sections.

5. Cross-linguistic mediation

One of the merits of the CEFR (2001) was that it introduced *mediation* as one of the four language activities and strategies alongside reception, production, and interaction (Council of Europe, 2001). Despite the novelty and the perceived importance of mediation as "written and/or oral activities [to] make communication possible between persons who are unable, for whatever reason, to communicate with each other directly" (Council of Europe, 2001, 14), the CEFR remained vague, not to say "rudimentary" (Reimann, 2019, 163), in defining this construct. Mediation activities, according to the CEFR (Council of Europe, 2001), were constrained to linguistic mediation of text, and pertained to activities such as "simultaneous interpretation", "consecutive interpretation", "informal interpretation", as well as "exact translation", "literary translation, "summarising gist", and "paraphrasing" predominantly across languages (ibid., 87). Apart from alluding to traditional translation and interpreting activities, the CEFR provided neither scales nor descriptors for mediation and thus left language practitioners, researchers, and test developers with a promising yet elusive construct. This has led to a variety of interpretations and operationalizations of the construct in testing contexts (e.g., Kolb, 2011; Piribauer & CEBS, 2016; Stathopoulou, 2015). As a result, it has even been argued that mediation might be a general educational ideal rather than a measurable competence (Reimann, 2020).

The CEFR-CV provides the long-awaited (sub)scales and descriptors for mediation activities (cf. Table 2), i.e., *Mediating a text*, *Mediating concepts* and *Mediating communication,* as well as for mediation strategies (cf. Table 3), i.e., *Strategies to explain a new concept* and *Strategies to simplify a text*. The focus is on the multidimensional nature of mediation (North & Piccardo, 2016), which considers language, in line with the action-orientation approach, as a "mediating tool" with a social, cultural, and cognitive function (North, 2021; North & Piccardo, 2019). In addition to mediating the content of a text, the CEFR-CV scales include aspects of literary competences and *Facilitating access to knowledge and concepts* by (co-)constructing meaning, especially in educational or collaborative contexts. The aim is to promote mutual understanding by overcoming individual, social, conceptual, or linguistic hurdles (Council of Europe, 2020, 91). Linguistic barriers may not only be different languages, but the CEFR-CV explicitly also mentions varieties, modalities and registers (ibid., 92). Overall, the new scales provide a fervent ground for language assessment in that they more clearly demonstrate "the cognitive and interpersonal challenge" inherent to mediation (North & Piccardo, 2016, 33) and link to other newly introduced scales, such as plurilingual/cultural competence. Hence the CEFR-CV provides language testers with a toolbox of illustrative descriptors for all competence levels, and raises new challenges as to how mediation scales may be exploited for assessment purposes, while encouraging reflection on earlier assessment practices (e.g., Reimann, 2019).

Mediation activities			
Mediating a text	Mediating concepts		Mediating communication
• *Relaying specific information* • *Explaining data* • *Processing text* • *Translating a written text* • *Note-taking* • *Expressing a personal response to creative texts* • *Analysis and criticism of creative texts*	**Collaboration in a group** • *Facilitating collaborative interaction with peers* • *Collaborating to construct meaning*	**Leading group work** • *Managing interaction* • *Encouraging conceptual talk*	• *Facilitating pluricultural space* • *Acting as an intermediary* • *Facilitating communication in delicate situations and disagreements*

Table 2: Mediation activities in the CEFR-CV (Council of Europe, 2020, 90)

Mediation strategies	
Strategies to explain a new concept	Strategies to simply a text
• *Linking to previous knowledge* • *Adapting language* • *Breaking down complicated information*	• *Amplifying a dense text* • *Streamlining a text*

Table 3: Mediation strategies in the CEFR-CV (Council of Europe, 2020, 90)

Although the CEFR-CV extends the mediation construct, cross-linguistic mediation is still key for communication in today's plurilingual societies and thus constitutes authentic use of language in many target-language situations. The new scales, needless to say illustrative rather than comprehensive, for *Mediating a text* such as *Relaying specific information* or *Translating a written text* in speech, sign or writing, and for mediating communication, such as *Facilitating pluricultural space* or *Acting as an intermediary in information situations*, may provide useful descriptors for test developers and expand the previous work on cross-linguistic tasks for assessment (e.g, Stathopoulou, 2015). In the Austrian context, Piribauer et al. (2016) provide noteworthy examples of how plurilingual oral

exams of mediation can be designed and implemented to assess different foreign languages.

It seems noteworthy, though, that the targeted construct, despite its additional specification by way of these new scales, might differ considerably depending on whether test takers are asked to complete an activity such as *Relaying information* in their first or any further language, and whether the interactive situation involves spoken, signed or written form. North (2021), for example, advises that for mediation tasks, it is necessary to give clear instructions on how test takers are expected to use their linguistic repertoire. Finally, test developers need to keep in mind that given the broadness of mediation in the CEFR-CV, these examples may be only one way of operationalizing the construct, yet there is not 'one' test of mediation. Instead, mediation tasks need to be adapted to local contexts and the communicative needs of a target test population, perhaps even more so than tests of the traditional skills (Lenz, 2022).

Beyond task development, cross-linguistic examinations also raise new challenges for rating performances. As a piloting of mediation tasks demonstrated, relational and conceptual aspects of mediation, while being fruitful for teaching, seemed to be problematic in formal testing contexts (North, 2021). Also, examiners need a clear understanding of the aspect of the mediation construct they are targeting and how it is related to and differs from similar competences, such as interaction or integrated skills (Kantarcıoğlu, 2022). Eventually, raters will face the question of equivalence when rating an examinee's performance as a mediator. While scales, such as *Acting as an intermediary in informal situations*, focus on mediating the "sense of what is said" (Council of Europe, 2020, 116), examiners will need to define what is a close enough translation of meaning and to what extent the message may be adapted for the recipient and the linguistic demand of the target situation. Consequently, not only mediation tasks but also rating scales and criteria for task fulfilment might need localization and adaptations for specific test populations and test uses.

6. Intralingual mediation

Apart from mediation across languages, the CEFR-CV also specifies and expands on the notion of *Mediating a text* as also including mediation activities that may occur in any language use scenario where the content of a particular text needs to be relayed to another speaker who does not have access to the particular text. Mediation in the CEFR-CV may therefore also be "intralingual", involving only one language, or different varieties or modalities of the same language, or even

different registers of the same variety. Put simply, it refers to any extraction of information from one text and the transmission of it in another (form of) communication. It thus touches on important areas such as *Processing text*, *Note-taking* or *Explaining data*, in other words the transformation of information found in graphs or diagrams into verbal text. For this, the CEFR-CV now offers illustrative scales to specify a proficiency progression that can then also be incorporated into curricula and assessments.

In trying to address many of the criticisms that the original 2001 framework had been confronted with, one of the areas of intralingual mediation that appears to have particular relevance for language education and assessment is the CEFR-CV's novel attention to the use of language as it relates to the reading, interpretation and analysis of literary texts. Teachers and teacher educators as well as literary scholars had often bemoaned that this area had been neglected, worrying about a marginalization of literary competences in language classrooms. The CEFR-CV is now offering three carefully calibrated scales to remediate this trend: 1) Reading as a leisure activity, 2) expressing a personal response to creative texts (including literature), and 3) analysis and criticism of creative texts (including literature). This naturally also has ramifications for CEFR-related language assessments at many levels as these additional scales further expand the assessment construct.

Since *Reading as a leisure activity* might be challenging to operationalize in a testing context, the focus in assessments may be on the two scales that describe the reactions or responses that literature often evokes as these may lend themselves better to implementation in assessment through speaking and writing tasks. The CEFR-CV makes a specific distinction between engagement/interpretation and analysis/evaluation, stating that the former, "[d]escribing a personal reaction and interpretation [,] is cognitively far simpler than giving a more intellectual analysis and/or evaluation" (Council of Europe, 2020, 106). *Expressing a personal response to creative texts (including literature)* therefore focuses on a learner explaining what they liked or what interested them about a literary text, describing characters and stating which they identified with, relating aspects of the creative text to their own experiences, feelings and emotions, and offering personal interpretation of the work as a whole or of aspects of it. *Analysis and criticism of creative texts (including literature)* is seen as intrinsically more challenging, which is why this form of engagement with literature is described to be more typical of secondary or tertiary educational settings, involving "more formal, intellectual reactions" (Council of Europe, 2020, 107). The scale concerns comparing different works of literature, giving a reasoned opinion of texts, and critically evaluating features of a piece (or pieces), including the effectiveness of

techniques employed. As such, the progression in this scale (in contrast to *Expressing a personal response*) is characterized by only starting at B1 level, and evaluative components only coming into play at B2 levels or above. These new scales on intralingual mediation are therefore further elements of the assessment construct, and if not considered adequately, any assessment claiming to be based on the CEFR may fall short of sampling representatively from the proficiency model.

Therefore, assessments may not only need to elicit a student's ability to "interpret and describe […] overall trends shown in simple diagrams" (Council of Europe, 2020, 97) or "take accurate notes in meetings and seminars on most matters likely to arise within their field of interest" (ibid., 105), but also their ability to "compare two works of literature, considering themes, characters and scenes, exploring similarities and contrasts and explaining the relevance of the connections between them" (ibid., 107) or "give his/her personal interpretation of the development of a plot, the characters and the themes in a story, novel, film or play" (ibid., 106). Again, these new CEFR-CV scales broaden the scope of what language functions, topics, or genres can and should be represented in assessment. However, it is not only the description of language use and communicative situations that are richer in the CEFR-CV compared to its predecessor, but also the individual levels themselves now provide more detailed descriptions of learners' target language use and proficiency.

7. Richer proficiency descriptions

The CEFR-CV provides richer descriptors for all language levels, allowing teachers as well as learners to make subtler distinctions between nuanced language skills. Particularly the assessment of (typically younger) learners with low-level language skills (pre-A1, A1 and A2) features its own intricacies which in the original CEFR had been insufficiently addressed. Generally, the literature in the language assessment field often neglects stages of early language learning, which is why guidelines on low-level assessment are scarce. The reasons might be grounded in the limited impact of low-level language tests compared to high stake tests that mainly target B1 or higher levels. However, the necessity for scales at low levels emerges from educational policies that aim to foster plurilingualism from an early age on (Alexiou & Stathopoulou, 2021).

One of the few studies that scrutinized the assessment of low-level language writing skills, for example, confirms the shortcomings of the original CEFR in terms of low-level descriptors by reporting that stakeholders' needs for the

assessment of basic language skills are not met and that there is a wide variety in approaches and expectations across test providers even within one lower proficiency level targeted (Konrad et al., 2018). The CEFR-CV took a first step in bridging this gap by implementing concrete descriptors for the lower levels and thereby highlight the importance of and need for early language learning in Europe (Alexiou & Stathopoulou, 2021). Thus, as one step to remedy this, the CEFR-CV introduces and defines a pre-A1 level as a stage in which "the learner has not yet acquired generative capacity, but relies upon a repertoire of words and formulaic expressions" (Council of Europe, 2020, 243).

Overall, 84 new scales have been added to the original CEFR, 46 of which target pre-A1 language learning situations that learners are likely to encounter in the real world. Alexiou and Stathopoulou (2021) provide an analysis of these scales with a focus on the richness and characteristics of pre-A1 level descriptors. They identified familiarity as the uniting feature of all pre-A1 descriptors as it is a prerequisite for low-level learners' ability to communicate and comprehend language input (Alexiou & Stathopoulou, 2021). Apart from that, brevity, support, slowness, clarity, and the repetition of individual words and formulaic expressions are further recurring elements of pre-A1 descriptors (Alexiou & Stathopoulou, 2021). The overall descriptor for oral comprehension at pre-A1 level, for example, requires the learner to be able to "recognize numbers, prices, dates and days of the week, provided they are delivered slowly and clearly in a defined, familiar everyday context". (Council of Europe, 2020, 48). Similarly, the descriptor for overall written production at pre-A1 says that learners "can give basic personal information (e.g. name, address, nationality), perhaps with the use of a dictionary" (ibid., 66). These pre-A1 can-do statements go hand in hand with the implementation of descriptors for young learners aged 7–10 and 11–15 years, which are available on the CEFR website. They have not been included in the CEFR-CV in order to avoid repetitive descriptors, as they are likely to overlap with pre-A1 descriptors. Even though the scales are far from being exhaustive, let alone complete, they are a good starting point for teaching and assessing young learners. They clarify what pre-A1 learners are required to know and be able to do at this level. Thus, it can be used as a tool for monitoring learners' progress and determine the level at which their language skills are and the input they need to further improve. It could also lay the foundation for developing more assessment tools for pre-A1 learners (Alexiou & Stathopoulou, 2021).

According to McElwee et al. (2019) the main potentials of the pre-A1 level descriptors are to render young learners' learning experience more coherent, strengthen the implementation of the action-oriented approach and to encourage plurilingual approaches. They also highlight the potential consequences on

assessment and its washback on teaching and learning in classrooms (McElwee et al., 2019). The new pre-A1 scales can be used to implement more systematic and transparent language teaching and assessment in primary schools or even pre-primary level, but also adults at low proficiency levels.

8. Conclusion

The new CEFR-CV to the CEFR does not constitute or offer a paradigm change. Rather, it enhances the paradigm shifting potential of the original CEFR, for it represents the logical next step in implementing the established action-oriented approach, revising key details and elaborating and enriching the original framework (Piccardo & North, 2019). This paper has attempted to exemplify six of these. The key changes related to (a) the native-speaker norm, (b) the phonological control scale, (c) increased digital communication, (d) interlingual and (e) intralingual mediation, and (f) the descriptions of lower-level learner competencies, have been discussed and some potential implications for language assessment have been outlined in light of the important assessment considerations of validity and authenticity. Like with the original CEFR, the CEFR-CV's innovations will likely take time to trickle down to the chalkface and into language assessments. De Jong (2022) for instance, laments that language test providers still conceptualize their assessments in a four skills, rather than a four modes approach. The CEFR-CV is therefore a renewed call to action for test developers to ensure action-oriented assessment instruments that represent a comprehensive spectrum of current, authentic language use, regarding its reference points and criteria, its communicative tasks, channels and modes, the range of learners/users and proficiency levels involved, and the increasing recognition of language users as mediators. Only in this way, assessments can be future-proofed and allow for valid score interpretations. The CEFR-CV is further a renewed call to action for SLA and testing researchers to investigate learner progressions along the proposed illustrative scales empirically. Above all, the CEFR-CV is testimony to the original framework's claim of being open, flexible and dynamic as it demonstrates the Council of Europe's willingness to respond to criticism and update the framework to changed and changing contexts (Bärenfänger et al, 2017). The onus is now on language teachers and test developers to follow suit.

References

Alderson, J. C. (2007). *The CEFR and the need for more research*. *The Modern Language Journal*, *91*(4), 659–663. https://doi.org/10.1111/j.1540-4781.2007.00627_4.x

Alexiou, T., & Stathopoulou, M. (2021). The pre-A1 level in the updated *Common European Framework of Reference for languages*. *Research papers in Language Teaching and Learning*, *11* (Special Issue: Innovations and Trends in Early Foreign Language Context), 11–29. https://tinyurl.com/3wkee9an (23.03.2023)

Bärenfänger, O., Harsch, C., Tesch, B., & Vogt, K. (2017). *Reform, Remake, Retousche? - Diskussionspapier der Deutschen Gesellschaft für Fremdsprachenforschung zum Companion to the CEFR*. https://dgff.de/assets/Uploads/Diskussionspapier-Companion-CEFR-Juni-2018.pdf (23.03.2023)

British Council, UKALTA, EALTA, & ALTE. (2022). *Aligning language education with the CEFR: A handbook*. https://tinyurl.com/mwtc9f86 (23.03.2023)

Bundesministerium für Bildung, Wissenschaft und Forschung (BMBWF) (w.d.). *Lebende Fremdsprachen*. Standardisierte Reife- und Diplomprüfung. https://www.matura.gv.at/srdp/lebende-fremdsprachen (23.03.2023)

Cinganotto, L. (2019). Online interaction in teaching and learning a foreign language: An Italian pilot project on the companion volume to the CEFR. *Journal of E-Learning and Knowledge Society*, *15*(1), 135–151. https://doi.org/10.20368/1971-8829/1618

Council of Europe (Ed.). (2001). *Common European Framework of Reference for languages: Learning, teaching, assessment.* Cambridge University Press. https://rm.coe.int/1680459f97

Council of Europe. (2020). *Common European Framework of Reference for languages: Learning, teaching, assessment—Companion volume*. Council of Europe Publishing. https://tinyurl.com/mwa24e5 (23.03.2023)

Davies, A. (2013). Native speaker. In C. A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics*. Blackwell Publishing. https://doi.org/10.1002/9781405198431.wbeal0855

De Jong, J. H. A. L. (2022). The action-oriented approach and language testing: A critical view. In D. Little & N. Figueras (Eds.), *Reflecting on the Common European framework of reference for languages and its companion volume* (70–80). Multilingual Matters.

Deygers, B. (2021). The CEFR companion volume: Between research-based policy and policy-based research. *Applied Linguistics*, *42*(1), 186–191. https://tinyurl.com/y57nmwb3

Deygers, B., Zeidler, B., Vilcu, D., & Carlsen, C. H. (2018). One framework to unite them all? Use of the CEFR in European university entrance policies. *Language Assessment Quarterly*, *15*(1), 3–15. https://doi.org/10.1080/15434303.2016.1261350

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

Galaczi, E., Post, B., Li, A., & Graham, C. (2011). Measuring L2 English phonological proficiency: Implications for language assessment. In J. Angouri, M. Daller, & J. Treffers-Daller (Eds.), *Proceedings of the 44th Annual Meeting of the British Association for Applied Linguistics, 1-3 Sept 2011*. Scitsiugnil Press.

Harding, L. (2011). *Accent and listening assessment: A validation study of the use of speakers with L2 accents on an academic English listening test*. Peter Lang.

Harding, L. (2016). What do raters need in a pronunciation scale? The users' view. In T. Isaacs & P. Trofimovich (Eds.), *Interfaces in second language pronunciation assessment: Interdisciplinary perspectives* (12–34). Multilingual Matters.

Harding, L. (2017). Validity in pronunciation assessment. In O. Kang & A. Ginther (Eds.), *Assessment in second language pronunciation* (30–48). Routledge.

Isaacs, T., & Harding, L. (2017). Pronunciation assessment. *Language Teaching*, *50*(3), 347–366. https://doi.org/10.1017/S0261444817000118

Isbell, D. R., & Kremmel, B. (2020). Test review: Current options in at-home language proficiency tests for making high-stakes decisions. *Language Testing*. https://doi.org/10.1177/0265532220943483

Kantarcıoğlu, E. (2022). The CEFR companion volume and mediation: An assessment perspective. In D. Little & N. Figueras (Eds.), *Reflecting on the Common European Framework of Reference for Languages and its companion volume* (182–193). Multilingual Matters.

Kolb, E. (2011). Wie stuft und prüft man Sprachmittlung? *Zeitschrift Für Fremdsprachenforschung*, *22*(2), 177–194. https://tinyurl.com/3j2yt25c (23.03.2023)

Konrad, E., Holzknecht, F., Schwarz, V., & Spöttl, C. (2018). *Assessing writing at lower levels: Research findings, task development locally and internationally, and the opportunities presented by the extended CEFR descriptors* (ARAGs Research Reports Online AR-G/2018/4). https://tinyurl.com/3xf4m3tv (23.03.2023)

Lenz, P. (2022). Some thoughts about the testing of mediation. In D. Little & N. Figueras (Eds.), *Reflecting on the Common European Framework of Reference for Languages and its companion volume* (129–138). Multilingual Matters.

Little, D. (2022). The Common European Framework of Reference for languages: Past, present and future: Introduction to part 1. In D. Little & N. Figueras (Eds.), *Reflecting on the Common European Framework of Reference for Languages and its companion volume* (20–27). Multilingual Matters.

Little, D., & Figueras, N. (Eds.) (2022). *Reflecting on the Common European Framework of Reference for languages and its companion volume*. Multilingual Matters.

McElwee, S., Devine, A., & Saville, N. (2019). Introducing CEFR pre-A1 descriptors for language instruction and assessment: Consequences, opportunities and responsibilities. *Zeitschrift Für Fremdsprachenforschung*, *30*(2), 215–232. https://tinyurl.com/434se84r (23.03.2023)

McNamara, T. (2014). 30 years on—Evolution or revolution? *Language Assessment Quarterly*, *11*(2), 226–232. https://doi.org/10.1080/15434303.2014.895830

McNamara, T., & Shohamy, E. (2016). Language testing and ELF: Making the connection. In M.-L. Pitzl & R. Osimk-Teasdale (Eds.), *English as a Lingua Franca: Perspectives and Prospects* (227–234). De Gruyter. https://doi.org/10.1515/9781501503177-030

Munro, M. J., & Derwing, T. M. (2011). The foundations of accent and intelligibility in pronunciation research. *Language Teaching*, *44*(3), 316–327. https://doi.org/10.1017/S0261444811000103

North, B. (2021). The CEFR companion volume—What's new and what might it imply for teaching/learning and for assessment? *CEFR Journal: Research and Practice*, *4*, 5–24. https://doi.org/10.37546/JALTSIG.CEFR

North, B. (2022). The CEFR companion volume project: What has been achieved. In D. Little & N. Figueras (Eds.), *Reflecting on the Common European framework of reference for languages and its companion volume* (51–64). Multilingual Matters.

North, B., & Piccardo, E. (2016). Developing illustrative descriptors of aspects of mediation for the Common European Framework of Reference (CEFR). A Council of Europe project. *Language Teaching*, *49*(3), 455–459. https://doi.org/10.1017/S0261444816000100

North, B., & Piccardo, E. (2019). Developing new CEFR descriptor scales and expanding the existing ones: Constructs, approaches and methodologies. *Zeitschrift Für Fremdsprachenforschung*, *30*(2), 143–161. https://www.dgff.de/assets/Uploads/ZFF-2-2019-01-North-Piccardo.pdf (28.03.2023)

Piccardo, E. (2016). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Phonological Scale Revision Process Report* (Council of Europe, Ed.). https://tinyurl.com/47hrafex (28.03.2023)

Piribauer, G., Atzelsberger, U., Greinix, I., Ladstätter, T., Mittendorfer, F., Renner, H., & Steinhuber, B. (2016). *Plurilingualism: Designing and implementing plurilingual oral exams. Framework for the Austrian upper secondary level oral leaving examination at colleges for higher vocational education exam subject plurilingualism*. https://tinyurl.com/yc2dtcd4 (28.03.2023)

Reimann, D. (2019). Mediation in den neuen Skalen und Deskriptoren des CEFR Companion Volume. *Zeitschrift Für Fremdsprachenforschung*, *30*(2), 163–180. https://www.dgff.de/assets/Uploads/ZFF-2-2019-02-Daniel-Reimann.pdf (28.03.2023)

Reimann, D. (2020). Sprachmittlung – Mediation im Fremdsprachenunterricht. Kompetenz und Bildungsziel. *Babylonia Journal of Language Education*, *3*(1), 10–21. https://doi.org/10.55393/babylonia.v3i1.36

Stathopoulou, M. (2015). *Cross-language mediation in foreign language teaching and testing*. Multilingual Matters. https://doi.org/10.21832/9781783094127

Widdowson, H. G. (1994). The ownership of English. *TESOL Quarterly*, *28*(2), 377–389. https://doi.org/10.2307/3587438