Developing new CEFR descriptor scales and expanding the existing ones: constructs, approaches and methodologies

Brian North[1] and Enrica Piccardo[2]

Der Beitrag gibt einen Überblick über die theoretischen Grundlagen des *Companion Volume with New Descriptors* zum "Gemeinsamen europäischen Referenzrahmen für Sprachen" und beschreibt die Entwicklung neuer Konzepte sowie die methodische Umsetzung ihrer empirischen Validierung. Dargestellt werden vor allem die Bereiche 'Mediation', 'Phonologie' und 'Gebärdensprache'. Weiterhin enthält der Companion Deskriptoren für junge Lernende auf dem Niveau "Pre-A1". Einige didaktische Konsequenzen dieser Entwicklungen werden ebenfalls erörtert.

1. The CEFR and its descriptors

The Common European Framework of Reference for Languages, Learning, Teaching, Assessment (CEFR; Council of Europe 2001) has been described as one of the most important curriculum documents of the last decades in relation to the integration of planning, enactment and evaluation (Graves 2008). The CEFR provides tools with which to work in a 'backwards design' (Richards 2013) by using the CEFR descriptors as a guide for the development of context-appropriate educational objectives (North 2014; North et al. 2018). Piccardo et al. (2011) provide teacher education materials in this regard and Piccardo (2014) explains the way that the CEFR concepts of the social agent and the action-oriented approach represent a significant development beyond the communicative approach. Indeed, the CEFR presents an innovative descriptive scheme that, among other things, goes beyond the traditional model of the four skills (listening, reading, speaking, writing) to propose four modes of communicative language activities and strategies: reception, production, interaction and mediation – the first two of which in spoken and written form make up the traditional four skills.

The CEFR provides scales of descriptors for different aspects of the descriptive scheme, which are intended to be used to define needs profiles for groups and individuals, plus proficiency profiles of what a person can currently do in a language. Examples of graphic profiles of both types are given in the

1   Korrespondenzadresse: Dr. Brian North, Eurocentres, Badenerstrasse 78, CH-8004 Zürich, Schweiz, E-Mail: bjnorth@eurocentres.com

2   Korrespondenzadresse: Dr. Enrica Piccardo, OISE/University of Toronto and Université Grenoble-Alpes, 6D-86 Gerrard Street East, M5B2J1 Toronto, Canada, E-Mail: enricapiccardo@gmail.com

CEFR Companion Volume (CV, Council of Europe 2018) that is the subject of this issue. The CEFR descriptors have had their critics, particularly language testers who would have preferred a systematic and exhaustive description of every aspect at every level (e.g. Alderson et al. 2006). A more justified criticism of the CEFR descriptors – made by researchers in second language acquisition – is that the progression suggested – though based on the use of these descriptors to assess the proficiency of real individuals – has not (yet) been fully confirmed by research on the development of proficiency over time (e.g. Alderson 2007). However, as Hulstijn (2007) pointed out, such research can be comfortably carried out whilst using the CEFR and as Hulstijn, Alderson and Schoonen (2010: 15) summarize: "It is fair to say that the resultant scales are probably the best researched scales of foreign language proficiency in the world." Issues to do with the CEFR descriptors are discussed in detail in North (2014) and more recently in a discussion paper prepared for the DGFF by Bärenfänger et al. (2018).

2. Background to the project

The CEFR was in many respects far ahead of its time, providing a basis for what has become known as a competence-based approach at a time when the latter development was in its infancy. However, there had been a number of significant developments since 2001, which led to a decision at a meeting in Strasbourg in May 2013 to launch a project to bring the CEFR up-to-date. For one thing, there had been further development of descriptors that could potentially help to fill out the original set, particularly as regards listening and reading, A1 and the C-levels. In particular, there were several projects that had developed and then calibrated descriptors to the CEFR scale of levels in ways comparable to those used in the original research. More fundamentally, the context of modern language teaching itself had developed considerably. Whereas in the late 1990s it was common in most European countries for classes to share the language of schooling as their mother tongue, by 2013 linguistic and cultural diversity had increased significantly in many countries. In addition, it was increasingly clear that students do not automatically acquire appropriate language to access knowledge, to formulate thoughts and (co-)construct meaning collaboratively. Learners who did not have the language of schooling as their mother tongue are thus further disadvantaged. In these circumstances, there was an increasing relevance of the CEFR's innovative view of the user/learner as a social agent and the related concepts of mediation and plurilingual and pluricultural competence, for which no CEFR descriptors had been provided in 2001. Furthermore, the paradigm change in language education suggested precisely by these innovative notions of

social agent, action-oriented approach, mediation and plurilingualism was reinforced by developments in theories related to language education since the 1990s when the CEFR was written, as explained in Section 5 below.

3. Structure of the project

The project took place in four separate and more or less consecutive stages:

1. Update of the 2001 scales, especially at the top (C levels) and bottom (addition of Pre-A1).
2. Development, validation and finalisation of new scales for areas not covered in 2001, notably: mediation including literature, plurilingual and pluricultural competence, and online interaction – plus a replacement of the scale for phonology.
3. Incorporation of descriptors for linguistic, pragmatic and sociolinguistic aspects of signing competence developed in a linked project at the Zurich University of Applied Science (ZHAW).
4. Collation of existing descriptors for younger learners for two age groups (7–10 and 11– 14), relating them to the expanded CEFR descriptors.

For reasons of space, following a brief outline of the first stage in the following section, this chapter will focus on the second stage, which became known as the mediation project. As can be appreciated, the context of the development was a very different one to the mid-1990s. At that time, following the intergovernmental Symposium held near Zürich that had recommended the CEFR and Portfolio, a Swiss National Research Programme project team developed and validated the set of levels and descriptors for the CEFR and a prototype portfolio (North 1995, 2000; North & Schneider 1998; Schneider & North 2000). The categories of the CEFR descriptive scheme were developed in an interactive process between the meetings of the CEFR Authoring Group[3] and the research in Switzerland. The idea behind the CEFR was that a common framework would lead on the one hand to increased transparency in the organisation of courses and the meaning of language qualifications and on the other hand to educational reform stimulated by the process of reflection and discussion beyond the boundaries of pedagogical cultures that the CEFR encourages.

Nowadays that Common Framework is a reality across Europe and beyond and many people have a detailed knowledge of the CEFR's educational

---

3   John Trim (coordinator), Daniel Coste, Brian North, Joe Sheils.

philosophy and share considerable expertise with regard to the CEFR levels and action-oriented approach. It was therefore possible to involve a large number of people in a principled manner in creating new CEFR descriptors. This was important in order to break the circularity of thinking that can occur in a small project group and consequently to base decisions on the inclusion/exclusion, formulation and level of descriptors on data and not just on the opinions of the developers.

The structure that the project adopted was the series of concentric circles typical of a Community of Practice (Wenger 2006). The main writing in each area was undertaken by a small Authoring Group[4] that worked interactively with a small Sounding Board of experts who provided sources, reacted by email, attended meetings and shared in decision-making about descriptors. In addition, there was a third tier of consultants for each area, with wider meetings with a group of some 20-30 consultants occurring at key stages in the project (July 2014, July 2015, and July 2016).

For the validation of the descriptors for mediation and related areas, described in Section 7 below, a fourth tier of coordinators were recruited to carry out workshops in a number of educational institutions (at first 140 and then around 190). These institutions were recruited through associations engaged with the CEFR. Finally, in a fifth tier, there were the team members in those institutions, and other individual project participants, over 1,500 in total. The coordinators and their team members were grouped into 'divisions' of approximately 30 institutions each: one from Eaquals (www.Eaquals.org), one from CercleS (www.cercles. org), one from Ealta (www.ealta.eu.org), one with German and American universities (particularly from UNIcert: www.unicert-online.org), and one international group. For the data collection in all three validation phases described in Section 7 below, care was taken to distribute materials evenly across these 'divisions' to prevent descriptors for any one category being investigated only in relation to one particular nationality, educational sector or type of institution.

4. Update of the 2001 scales

The main aim of the update was to provide more descriptors for reception, which had been under-represented in 2001, and to give more detail at the top and bottom of the scales. The main sources of descriptors for this purpose were a series of projects that had also used a principled methodology to develop descriptors and then to validate and calibrate them mathematically to CEFR levels using the

---

4   Brian North (coordinator), Tim Goodier, Enrica Piccardo, Maria Stathopoulou.

Rasch scaling model (Linacre 2015). These projects were: ALTE Can Do-Statements (particularly useful for the domains of work and study), AMKKIA (a Finnish project for aspects of linguistic competence), the 'common scales' from Cambridge English Assessment (for aspects of pragmatic competence), CEFR-J (especially descriptors for very low levels, produced in relation to teaching English in secondary school in Japan; Negishi, Takada & Tono 2013), English Profile (for the C levels; Green 2012), Lingualevel (produced for lower secondary in Switzerland; Lenz & Studer 2007), and Pearson's Global Scale of English (developed from CEFR descriptors; De Jong, Mayer & Hayes 2016). In addition to filling out existing scales, especially at the top and bottom, these sources were also used to provide scales for 'Reading as a leisure activity', 'Using telecommunications', and 'Sustained monologue: Giving information'.

The descriptors selected were edited into the CEFR 'house style' when necessary, and then went through an extensive process of consultation with experts over a period of some 15 months. There are no changes at all to CEFR levels – only the provision of more detail. Some changes of formulation to 16 existing descriptors (listed in Appendix 7 of the Companion Volume; Council of Europe 2018) have been made, as explained below, but no changes have been made to the levels themselves.

4.1 Changes to 2001 formulation

There were two main issues involved here: First, the removal of 'absolute' terminology in certain uncalibrated C2 descriptors. C2, as the CEFR says, "is not intended to imply native speaker or near native speaker competence. What is intended is to characterise the degree of precision, appropriateness and ease with the language which typifies the speech of those who have been highly successful learners" (Council of Europe 2001: 36). Not all of the C2 descriptors were calibrated in the Swiss project, and this led to a few cases of exaggerated wording, which is now corrected. For example, the C2 descriptor for 'Overall listening comprehension' has been reformulated as follows, with the part in italics replacing the part struck through: "*Can understand with ease virtually* ~~Has no difficulty with~~ any kind of spoken language, whether live or broadcast, delivered at fast ~~native~~ *natural* speed". Most of the changes, however, concerned removing the expression 'native speaker', since this notion has been increasingly criticised since the early 1990s (e.g. Kramsch 1993; Byram & Zarate 1996). Apart from 'native speed' in the above descriptor, there were only three other occurrences at C2. The bulk of the descriptors affected concerned two concepts that reoccur in different

scales at B2/B2+: the ability to keep up with an animated conversation and the ability to maintain relationships without other people having to accommodate their behaviour to the user/learner. In neither case was there any intention to compare the user/learner's competence and that of a 'native speaker' – and in neither case was the stress on seeing conversational partners as 'native speakers': the issues are fast colloquial speech and assumptions of familiarity with the cultures of a community, not whether the language used by speaker is their native language. These changes and clarifications were overwhelmingly approved in the consultation with both individuals and institutions which preceded publication.

4.2 Pre-A1 level

The CEFR-J and Lingualevel projects provided the main source for the new Pre-A1 level. They had each developed and calibrated descriptors to this band of proficiency approximately halfway to A1, at which users/learners deploy learnt phrases, supported by gesture. The existence of such proficiency approximately halfway towards A1 is discussed in CEFR Section 3.6, which suggests that "[i]n certain contexts, for example with young learners, it may be appropriate to elaborate such a 'milestone'" (Council of Europe 2001: 31). In the CEFR this milestone was referred to as 'Tourist', because of the nature of the six descriptors calibrated there in the Swiss project. Since there were only six descriptors, they were included in Level A1 in 2001, as they form an early part of that level.

4.3 Phonology

The most significant change to the 2001 descriptors is the complete replacement of the holistic scale for phonological control with an analytical one for (a) overall phonological control, (b) sound articulation, and (c) prosody (stress, rhythm and intonation). The new scale focuses on intelligibility as the crucial factor and acknowledges that accent may remain even at C2. The original phonology scale had been the least successful scale in the research project behind the 2001 descriptors (see North 2000: 238-241). In particular, it contained the following descriptor for B2 (actually calibrated a high B2+): "Has acquired a clear, natural, pronunciation and intonation." Apart from being an unrealistic aim for B2, 'natural' appeared to be an echo of the 'native speaker' that seemed inappropriate in the context of a project whose main focus, as described in the rest of this paper, was the development of descriptors for mediation and plurilingual/pluricultural

competence. A full report on the phonology project is available (Piccardo 2016). The project followed the same development, validation and calibration procedures as the mediation project (see Section 7).

4.4 Literature and online interaction

The Council of Europe's Educational Policy Department reported that CEFR descriptors had frequently been requested for these two areas, so it was a priority to include them in the project. The receptive aspect of reading literature is covered in the new scale for 'Reading as a leisure activity' mentioned above. Creative texts, particularly literature and films, provoke discussion that itself has an important function in intercultural education, particularly at secondary level. Literature is valued in education because it mediates between the inner and outer worlds, broadens perspective and gives a window on other experiences, other cultures. Two scales for reactions to creative text were therefore developed during the mediation project discussed below. The main inspiration for the descriptors came from classic texts concerning the teaching and assessment of literature at secondary level (e.g. Oatley 1994; Purves 1971) plus descriptors included in portfolios. A distinction is made between two scales for (a) engaging with the work on a personal level and (b) interpreting and evaluating it in its context on the other. The former scale has more focus on lower levels, the latter on higher levels.

The CEFR makes a distinction between written interaction and written production, which many language professionals at the time considered to be exaggerated. As a result, the self-assessment grid for the portfolio passports was, perhaps unfortunately, simplified by reuniting these into 'writing'. For the sake of coherence, it was then decided to replace the existing CEFR self-assessment grid (CEFR Table 2) with this simplified version in the 2001 publication. In retrospect, this early insistence on a fundamental distinction between coherent, drafted and redrafted written production on the one hand and spontaneous written chat on the other (see e.g. Halliday 1989) has been vindicated by the subsequent development of email and online chat. Since these digital developments involve a different kind of communication, with greater need for clarity and redundancy and the integration of different media, it was decided to develop two scales for online interaction, rather than revisiting written interaction. One scale focuses on the interpersonal aspects of conversation and discussion, the other on transactional aspects of internet use – from booking flights to coordinating online collaborative work. In this case, there was little literature available to inspire descriptors – apart

from on the issue of integrating media – so the constructs for the scales were developed from scratch by Enrica Piccardo and Tim Goodier, the two members of the team who had expertise in this area.

In the Companion Volume (Council of Europe 2018), each scale, new and old, has a short introduction that gives the rationale, the concepts involved and the way progression up the levels operates.

5. Conceptualisation of mediation

The most complex process of conceptualisation concerned mediation itself (for a detailed explanation of the reasoning behind the scheme, as well as of the development and validation of the descriptors see North & Piccardo 2016). The aspect of mediation most apparent in the CEFR is '(cross-)linguistic mediation'. This is by no means confined to professional interpretation and translation. It is a relatively common occurrence to be asked to give an idea of the main content of a text that another person cannot understand, or to "act as an intermediary between interlocutors who are unable to understand each other directly, normally (but not exclusively) speakers of different languages" (Council of Europe 2001 and 2018: 87). As Backus et al. (2013) point out, studies have shown that ordinary people are fully capable of performing these functions in informal situations, and they have begun to appear in content specifications (Glaboniat et al. 2005) as well as examinations, both oral (Piribauer et al. 2015) and written (Stathopoulou 2013). Here, therefore, there was a wealth of source material to start from.

However, Piccardo (2012) shows that mediation as presented in the CEFR also implies social and cultural mediation. On this theme, Zarate (2003) develops Kramsch's (1993) notion of 'third space' as an "alternative to linguistic and cultural confrontation. In this plural area difference is pinpointed, negotiated and adapted" (Zarate 2003: 95) by focusing on the relations involved in 'in-between spaces,' thus developing 'symbolic competence' (Kramsch 2009).

Piccardo (2012) also points out that the CEFR gives conceptual mediation – in the sense in which it is used in relation to concept formation in the sociocultural approach (Lantolf 2000; Lantolf & Poehner 2014) that has been developed from Vygotsky's views – a central place in its scheme to describe language learning and use.

> Vygotsky repeatedly emphasized the role of mediation in the development of reflexive self-determining human agency, or "active adaptation" (Vygotsky 1981: 151-152). Humans internalized their own evolution while securing change in their environment, remaking both their conditions of existence and themselves. (Marginson & Dang 2017: 119).

In addition, the CEFR concept of the user/learner as social agent and its emphasis on the interaction between the social and individual in relation to both the user/learner's internal competences and mental context and the external context of domain and situation, mirror key concepts informing an ecological perspective (van Lier 2004, 2010), in addition to sociocultural theory. From this perspective, learning occurs through "perception in action" (van Lier 2004: 97) when the social agent's attention fixes on 'affordances'. An affordance is an "opportunit[y] for action in the environment" (Käufer & Chemero 2015: 166), something which the user/learner needs in order to accomplish a task. Van Lier sees 'meaning potential' (Halliday 1973) as "the semiotic potential or the affordances" (van Lier 2004: 74) offered. Social agents act on affordances when they exercise their agency by accepting invitations that they perceive in the environment. Mediation (for oneself or from a third party) helps people to see the affordances. But when a social agent acts, they interact with the actions of others and with the environment itself in complex ways that lead to emergence of new phenomena and states and to – often unpredictable – change over time, as predicated in complexity theories (Larsen-Freeman 1997; Larson-Freeman & Cameron 2008). The way in which the conceptualisation of mediation in the Companion Volume (Council of Europe 2018) echoes social cognitive theories of agency (Bandura 1989, 2001), sociocultural theory and the theory of affordances as well as complexity theories and approaches to collaborative learning is explained in more detail in Piccardo and North (2019).

An essential aspect of mediation, in relation to the recognition of affordances and the development of concepts, is the use of language as a tool in order to articulate thought. The process of thinking something through alone or in discussion has been called 'languaging' (Swain 2006) or 'plurilanguaging': "a dynamic, never-ending process to make meaning using different linguistic and semiotic resources" (Piccardo 2017: 9). This may well be for the primary benefit of others rather than for oneself, helping them to understand something. This will include creating the conditions that facilitate communication and understanding as well as such conceptual mediation itself. When teachers provide scaffolding to help learners to understand something, or explain something in one language and then support it with another, when a teacher or a group member encourages someone to follow through their own thoughts, this is also mediation. This type of mediation occurs inside groups in all collaborative group work in academic and professional life – even more so in an online environment. Often in such contexts there may be a cross-linguistic and cross-cultural element as part of the pooling of knowledge and resources, which requires both plurilingual and pluricultural competences on the part of the participants.

The complete list of new scales created for mediation and its related areas is as shown in Table 1. The scales are presented in four groups: for mediating texts, mediating concepts, mediating communication, and mediation strategies, plus two new scales for online interaction, the one for phonological control, and three for plurilingual and pluricultural competence. Titles in italics are the names of groups; titles in plain text are scales.

Table 1: New descriptor scales

Overall mediation
Mediating texts
Relaying specific information in speech
Relaying specific information in writing
Explaining data (e.g. in graphs, diagrams, etc.) in speech
Explaining data (e.g. in graphs, diagrams, etc.) in writing
Processing text in speech
Processing text in writing
Translating a written text in speech
Translating a written text in writing
Note-taking (lectures, seminars, meetings, etc.)
Expressing a personal response to creative texts (including literature)
Analysis and criticism of creative texts (including literature)
Mediating concepts
Facilitating collaborative interaction with peers
Collaborating to construct meaning
Managing interaction
Encouraging conceptual talk
Mediating communication
Facilitating pluricultural space
Acting as intermediary in informal situations (with friends and colleagues)
Facilitating communication in delicate situations and disagreements
Mediation strategies
Linking to previous knowledge
Adapting language
Breaking down complicated information
Amplifying a dense text
Streamlining a text
Online interaction
Online conversation and discussion
Goal-oriented online transactions and collaboration
Communicative language competence
Phonological control
Plurilingual and pluricultural competence
Building on pluricultural repertoire
Plurilingual comprehension
Building on plurilingual repertoire

6. Development approach

The development and validation of the descriptors for mediation and related areas took place over six stages between the beginning of 2014 and the end of 2017:

1. Development   (January 2014 – February 2015)
2. Validation   (February 2015 – February 2016)
3. Analysis, revision   (February – May 2016)
4. Consultation   (June 2016 – February 2017)
5. Piloting   (January – July 2017)
6. Finalisation   (August – October 2017)

The place to start in developing descriptors for a new area is not the CEFR levels or their existing descriptors, but rather the kind of language behaviour that is valuable as an educational objective in the area(s) under study. The project followed a design-based research philosophy (van den Akker et al. 2006) in an iterative process. The steps in producing draft descriptors to then be validated were basically the following:

(a) collect concepts and behaviours from the literature and from reflection and discussion;
(b) collect existing descriptors that may be relevant, even if not validated;
(c) define the main concepts and behaviours in the area concerned;
(d) formulate these into early, draft descriptors;
(e) consult small-scale with informants to select the better descriptors and to improve formulation;
(f) consult small-scale with informants to see what level the descriptors seem to be;
(g) re-analyse what the construct(s) in the emerging scale appear(s) to be, and how they appear to develop up the levels;
(h) formulate more descriptors to 'plug gaps' in the draft grid of construct(s) and levels;
(i) repeat steps (e) to (h) until satisfied.

Step (i) was completed by March 2015. The Authoring Group had worked in different combinations in sub-groups focusing on different areas, discussing drafts with each other, with key consultants in the area concerned and with the Sounding Board, who had attended a series of meetings between September 2014 and February 2015. Most development projects in education actually jump from

such a development stage straight to consultation and/or piloting before publication, without any formal validation. By contrast, the approach to development and validation taken in the mediation project was a further development of the methodologies of the 1993-6 Swiss National Research Programme project (North 1995, 2000; North & Schneider, 1998, Schneider & North 2000).

By March 2015, 427 new descriptors were ready to enter the validation process. This was not, however, the end of the development. After each phase of validation, descriptors were dropped, amended and added following the iterative process of steps (e) to (h). The descriptors that came out of the entire validation process in December 2015 were far shorter, more focused and better covered the full range of CEFR levels than had been the case with the 427 descriptors from March 2015. Approximately 35% had been replaced or eliminated and over half of the remainder had been substantially revised as a result of comments. However, perhaps because of the very nature of mediation, there was a truly excessive number of descriptors for Level B2 – and especially B2+. Therefore, finalisation involved a pruning process. Over one hundred descriptors that were validated, but then dropped at this stage, are included as 'supplementary descriptors' in the final appendix of the Companion Volume (Council of Europe 2018).

Consultation on the result of the project included a meeting of experts and a wider survey of some 60 experts, followed by formal consultation of member states, key institutions and associations, plus individuals (of whom 500 completed the questionnaire), with written submissions by various associations, and a series of meetings with a 'follow up group.' All but two of the scales (for mediation strategies) were considered helpful by 80% or more of both institutional and individual respondents. The most popular scales among member states were those for plurilingual and pluricultural competence.

7. Validation

This section describes briefly the validation and analysis processes (Stages 2 and 3 of the six listed at the beginning of the previous section). These were very similar to those undertaken for the 2001 descriptors, but on a far larger scale. The first two main validation phases replicated face-to-face workshops undertaken in the Swiss project; the third phase was similar to the main data collection in that project, collecting data to calibrate the descriptors with the Rasch model. A more detailed description of the validation is available in North and Piccardo (2016). As mentioned in Section 3 above, by February 2015, 140 institutions had by

February 2015 been recruited through Ealta, Eaquals, CercleS, UNIcert and other associations for the validation process. This rose to 189 institutes by the second phase. For each of the three main validation phases, a systematic distribution of a set of 20-30 overlapping questionnaires across the different 'divisions' of participating institutions ensured a reasonably even sampling. The process itself followed a mixed methods, sequential qualitative and quantitative approach (Creswell & Plano Clark 2011). The three main phases were themselves followed by two supplementary phases with volunteers: one to provide further validation of the descriptors for plurilingual and pluricultural competence, and the other for the phonology sub-project.

Validation phase 1 consisted of a series of 140 face-to-face workshops (with approximately 1000 participants in total) organised in the participating institutions in March 2015. Participants worked in pairs on paper. The tasks were (a) to assign a list of descriptors presented in alphabetical order to related but distinct categories – with an opportunity to suggest dropping descriptors, (b) to evaluate each descriptor for clarity, pedagogic usefulness and relevance to real world language use, and (c) to propose improvements to the formulation. It was here that informants struck out subordinate clauses and, in general, helped to tighten up descriptor wording.

Validation phase 2 consisted of another 189 face-to-face workshops with approximately 1300 participants in total organised by the institutional coordinators in May-June 2015. This time the focus was on assigning descriptors to CEFR levels. First participants undertook familiarisation activities as recommended in the manual for relating assessments to the CEFR (Council of Europe 2009). Then, again working on paper in pairs, participants discussed the level of the descriptors. Finally, they entered their considered, individual judgement on the level of the descriptors into a SurveyMonkey online.

Validation phase 3 consisted of an online survey with approximately 3500 usable responses (75% in English, 25% in French), which was the main data collection for calibration. The task was to assess the language ability of a person one knew well (which could be oneself), saying to what extent that person could do what was described in the descriptor. Participants were asked to use a defined 0-4 rating scales to do so. This 0-4 rating scale was a replication of the assessment task used to provide the data to scale the original descriptors and create the CEFR levels (North 1995, 2000; North & Schneider 1998).

As mentioned in the previous section, comments made in each of the three phases, as well as in the formal consultation, were taken into account in the final version. In validation phase one, there were hundreds of suggestions for reformulations and comments. Validation phases two and three together produced

just over 1,100 comments, with another 2,900 comments received – many quite lengthy – during the consultation.

Validation phases 2 and 3 provided data to calibrate the new descriptors to the existing CEFR scale of levels with the help of a Rasch model analysis. This is done by 'anchoring' the scale produced in 2016 back to that behind the 2001 descriptors by reusing some existing CEFR descriptors and then double-checking that the new descriptors fitted into the old scale. This is a slightly simplified account, since in the original research, descriptors for reading needed to be separately anchored to the CEFR scale, using the listening descriptors as 'anchors' (see North 2000 and Schneider & North 2000). This 'anchoring' method, common in Rasch modelling, had also been successfully used in several of the projects that were sources to update the 2001 scales. The process is best documented by de Jong, Mayer and Hayes (2016). The process in the current project was slightly more complicated in practice, with the final Rasch 'difficulty values' for the descriptors being calculated in separate analyses for each of the main groups of categories (see North & Piccardo 2016 for details).

For the phonology scale, development followed exactly the same pattern as in the main project, with some 250 participants as informants (see Piccardo 2016; Piccardo & North 2017). In this case, validation phase 3 took the form of evaluating performance on video clips, using the same 0-4 rating scale from the Swiss project that had been used in validation phase 3 of the main project. The resultant scale was linked back to the scale behind the 2001 descriptors through two standard-setting methods: (a) 'anchoring' descriptors adapted from Cambridge English Assessment scales with estimated difficulty values, and (b) a simplified version of the 'bookmark method', recommended in the manual for relating assessments to the CEFR (Council of Europe 2009).

The sign language project carried out by the Zurich University of Applied Science (ZHAW) in another Swiss National Research Programme project followed a similar process to the mediation project, except that (a) it was completely data-based and (b) everything took place on a smaller scale since the signing community is a small one. In a first phase, expert signers were filmed performing different types of texts. Then sign language teachers discussed the competences they saw in these performances and the Authoring Group[5] formulated those into draft descriptors. The sign language descriptors were then assigned to categories in workshops, and an attempt made to fill out a grid of levels and categories, drafting descriptors accordingly. Then the descriptors were scaled for difficulty in an online survey, again using the Rasch model to analyse responses on a simplified 0-4 rating scale for each descriptor. The final step,

---

5   Jörg Keller, Petrea Bürgin, Aline Meili, Dawei Ni.

setting the standards for the CEFR levels, was undertaken with the adapted version of the bookmark method that had also been used for phonology.

8. The product: The CEFR Companion Volume

The product from the project is presented in the CEFR Companion Volume (Council of Europe 2018). The Companion Volume is more than just a set of descriptor scales; it contains the following:

- A text on key aspects of the CEFR for teaching and learning languages.
- The complete set of CEFR illustrative descriptors, old and new, with a graphic overview of each section and with a short rationale for each scale.
- Overviews of the CEFR levels in Appendices 1-4:
    - an abridged version of CEFR Section 3.6 explaining the salient characteristics of the levels;
    - the CEFR self-assessment grid (CEFR Table 2) updated to include mediation;
    - the grid showing qualitative aspects of performance (CEFR Table 3) updated to include phonology, and
    - a writing assessment grid taken from the CEFR manual for relating examinations to the CEFR (Appendix 4).
- Examples extending the descriptors on the new scales for the personal, public, occupational and educational domains (Appendix 5).

The text on key aspects for teaching and learning clarifies the CEFR descriptive scheme, the way the levels and descriptors are intended to be used, and the paradigm shift in language education that is implied by the CEFR notion of the user/learner as a plurilingual/pluricultural social agent. It also shows the relevance of the CEFR beyond foreign languages, emphasising the use of language as a tool for articulating thought and as a vehicle for facilitating understanding, rather than just as a way to communicate a message. It provides straightforward texts and simple visuals that may be useful in teacher education to help get across the multidimensional, action-oriented approach of the CEFR, and the complex (as opposed to linear, simple) vision of language education that it promotes.

A series of 68 pilots were carried out in 2017 and the preliminary version of the Companion Volume was put online in English in late October. The full version, including scales for different aspects of signing competence, was

published online in English and French in January 2018, with the Council of Europe ISBN publications planned to follow at the end of 2019.

A conference entitled "Building Inclusive Societies through Enriching Plurilingual and Pluricultural Education at a Grassroots Level: the Role of the CEFR Companion Volume" took place in Strasbourg in May 2018 to launch the Companion Volume. PowerPoints and videos from a webinar in January 2018 are available on the Council of Europe website. There are plans to make available an online descriptor search tool with which users will be able to select the language, categories and levels of descriptors for download. During the current academic year 2018-2019, some 30-40 'official' case studies of implementing the Companion Volume are taking place, with the intention of publishing a volume recording these experiences and the relevance of the new descriptors to different areas of language education.

This paper has tried to provide insight into the thorough research process that informed the Companion Volume (Council of Europe 2018), which builds on the methodology originally used for the CEFR. The Companion Volume represents a major milestone in language education and the wealth of initiatives that are being undertaken around it can have a leverage effect to spark innovation in the field. These various initiatives, including this current publication, are intended to encourage readers to investigate and experiment with the new and updated descriptors in relation to curriculum planning, task design and teacher assessment, and so to foster inclusive, action-oriented, plurilingual and intercultural language education.

<div align="right">Eingang des revidierten Manuskripts 08.04.2019</div>

References

Alderson, J. Charles (2007): The CEFR and the Need for More Research. *The Modern Language Journal* 91: 4, 659-662.

Alderson, J. Charles; Figueras, Neus; Kuijper, Henk; Nold, Günter; Takala, Sauli & Tardieu, Claire (2006): Analysing Tests of Reading and Listening in Relation to the Common European Framework of Reference: the experience of the Dutch CEFR Construct Project. *Language Assessment Quarterly* 3: 1, 3-30.

Backus, Ad; Gorter, Durk; Knapp, Karlfried; Schjerve-Rindler, Rosita; Swanenberg, Jos; ten Thije, Jan D. & Vetter, Eva (2013): Inclusive Multilingualism: Concept, Modes and Implications. *European Journal of Applied Linguistics* 1: 2, 179-215.

Bandura, Albert (1989): Human Agency in Social Cognitive Theory. *American Psychologist* 44, 1175-1184.

Bandura, Albert (2001): Social Cognitive Theory: An agentic perspective. *Annual Review of Psychology* 52, 1-26.

Bärenfänger, Olaf; Harsch, Claudia; Tesch, Bernd & Vogt, Karin (2018): Reform, Remake, Retousche? Diskussionspapier der Deutschen Gesellschaft für Fremdsprachenforschung zum *Companion to the CEFR (2017)* [Online: http://dgff.de/assets/Uploads/ Diskussionspapier-Companion-CEFR-Juni-2018.pdf, 09.03.2019].

Byram, Michael & Zarate, Geneviève (1996): Defining and Assessing Intercultural Competence: Some principles and proposals for the European context. *Language Teaching* 29: 4, 239-243.

Council of Europe (2001): *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press [Online: https://www.coe.int/en/web/common-european-framework-reference-languages, 18.08.2018].

Council of Europe (2009): *Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR)*. Strasbourg: Council of Europe. [Online: www.coe.int/en/web/common-european-framework-reference-languages/tests-and-examinations, 18.08.2018].

Council of Europe (2018): *Common European Framework of Reference for Languages: Learning, Teaching, Assessment – Companion Volume with New Descriptors*. Strasbourg: Council of Europe Publishing [Online: www.coe.int/en/web/common-european-framework-reference-languages, 18.08.2018].

Creswell, John W. & Plano Clark, Vicky L. (2011): *Designing and Conducting Mixed Methods Research* (2nd ed.). Thousand Oaks, CA: Sage.

De Jong, John H.A.L; Mayer, Mike & Hayes, Catherine (2016): *Developing Global Scale of English Learning Objectives aligned to the Common European Framework*. Global Scale of English Research Series. Pearson [Online: www.pearsonelt.com/content/dam/professional/ english/pearsonelt.com/TeacherResources/GSE/GSE-WhitePaper-Developing-LOs.pdf, 18.08.2018]

Glaboniat, Manuela; Müller, Martin; Rusch, Paul; Schmitz, Helen & Wertenschlag, Lukas (2005): *Profile Deutsch A1 - C2. Lernzielbestimmungen, Kannbeschreibungen, Kommunikative Mittel*. München: Langenscheidt.

Graves, Katherine (2008): The Language Curriculum: A social contextual perspective. *Language Teaching* 41: 2, 147-181.

Green, Antony (2012): *Language Functions Revisited: Theoretical and empirical bases for language construct definition across the ability range*. Cambridge, UCLES/Cambridge University Press.

Halliday, Michael A.K. (1989): *Spoken and Written Language*. Oxford: Oxford University Press.

Hulstijn, Jan H. (2007): The Shaky Ground Beneath the CEFR: Quantitative and qualitative dimensions of language. *Modern Language Journal* 91: 4, 663-666.

Hulstijn, Jan H.; Alderson, J. Charles A. & Schoonen, Rob (2010): Developmental Stages in Second-language Acquisition and Levels of Second-language Proficiency: Are there links between them? In: Bartning, Inge; Martin, Maisa & Vedder, Ineke (eds.): *Communicative Proficiency and Linguistic Development: intersections between SLA and language testing research*, 11-20 [Online: www.eurosla.org/publications/eurosla-monographs-discontinued-from-2017/, 18.08.2018].

Käufer, Stefan & Chemero, Anthony (2015): *Phenomenology: An Introduction*. Cambridge: Polity Press.

Kramsch, Claire (1993): *Context and Culture in Language Teaching*. Oxford: Oxford University Press.

Kramsch, Claire (2009): *The Multilingual Subject*. *What Language Learners Say About Their Experience and Why It Matters*. Oxford: Oxford University Press.

Lantolf, James P. (ed.) (2000): *Sociocultural Theory and Second Language Learning*. Oxford: Oxford University Press.

Lantolf, James P. & Poehner, Matthew E. (2014): *Sociocultural Theory and the Pedagogical Imperative in L2 Education: Vygotskian Praxis and the Research/practice Divide*. New York: Routledge.

Larsen-Freeman, Diane (1997): Chaos/complexity Science and Second Language Acquisition. *Applied Linguistics* 18: 2, 141-165.

Larsen-Freeman, Diane & Cameron, Lynne (2008): *Complex Systems and Applied Linguistics*. Oxford: Oxford University Press.

Lenz, Peter & Studer, Thomas (2007): *Lingualevel*. *Instrumente zur Evaluation von Fremdsprachenkompetenzen 5. bis 9. Schuljahr*. Bern: Schulverlag [Online: www.lingualevel.ch, 18.08.2018].

Linacre, J. Mike (2015): *Winsteps: Rasch-model Computer Program*, Chicago: MESA Press.

Marginson, Simon & Dang, Thi Kim Anh (2017): Vygotsky's Sociocultural Theory in the Context of Globalization. *Asia Pacific Journal of Education* 37: 1, 116-129 [Online: http://dx.doi.org/10.1080/02188791.2016.1216827, 18.08.2018].

Negishi, Masashi; Takada, Tomoko & Tono, Yukio (2013): A Progress Report on the Development of the CEFR-J. In: Galaczi, Evelina D. & Weir, Cyril J. (2013) (eds.) *Exploring Language Frameworks: Proceedings of the ALTE Krakow Conference, July 201*. Cambridge: Cambridge University Press, 135-163.

North, Brian (1995): The Development of a Common Framework Scale of Descriptors of Language Proficiency Based on a Theory of Measurement. *System* 23: 4, 445-465.

North, Brian (2000): *The Development of a Common Framework Scale of Language Proficiency*. New York: Peter Lang.

North, Brian (2014): *The CEFR in Practice*. Cambridge: Cambridge University Press.

North, Brian; Angelova, Mila; Jarocsz, Elizabeta & Rossner Richard (2018): *Language Course Planning*. Oxford: Oxford University Press.

North, Brian & Piccardo, Enrica (2016): *Developing Illustrative Descriptors of Aspects of Mediation for the Common European Framework of Reference (CEFR)*. Strasbourg, France: Council of Europe [Online: www.coe.int/en/web/common-european-framework-reference-languages/documents, 18.08.2018].

North, Brian & Schneider, Günther (1998): Scaling Descriptors for Language Proficiency Scales. *Language Testing* 15: 2, 217-262.

Oatley, Keith (1994): A Taxonomy of the Emotions of Literary Response and a Theory of Identification in Fictional Narrative. *Poetics* 23, 53-74.

Piccardo, Enrica (2012): Médiation et apprentissage des langues: Pourquoi est-il temps de réflechir à cette notion ? [Mediation and Learning Languages: Why is it time to reflect on this notion?] *ELA: Études de Linguistique Appliquée [Studies in Applied Linguistics]* 167, 285-297.

Piccardo, Enrica (2014): *From Communicative to Action-oriented: a Research Pathway* [Online: https://transformingfsl.ca/en/resources/from-communicative-to-action-oriented-illuminating-the-approaches/, 19.06.2019].

Piccardo, Enrica (2016): *Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Phonological scale revision process. Report*. Strasbourg: Council of Europe [Online: https://rm.coe.int/phonological-scale-revision-process-report-cefr/168073fff9, 18.08.2018].

Piccardo Enrica (2017): Plurilingualism as a Catalyst for Creativity in Superdiverse Societies: A systemic analysis. *Frontiers*. *Psychology* 8: 2169 [DOI: 10.3389/fpsyg.2017.02169].

Piccardo, Enrica; Berchoud, Marie; Cignatta, Tiziana; Memntz, Olivier & Pamula, Malgorzata (2011): *Pathways Through Assessment, Learning and Teaching in the CEFR*. Graz, Austria: European Centre for Modern Languages.

Piccardo, Enrica & North, Brian (2019): *The Action-oriented Approach: a Dynamic Vision of Language Education*. Bristol: Multilingual Matters.

Piribauer, Gerda; Atzlesberger, Ute; Greinix, Irmgard; Ladstätter, Thomas; Mittendorfer, Franz; Renner, Helmut & Steinhuber, Belinda (2015): *Plurilingualism: Designing and Implementing Plurilingual Oral Exams: Framework for the Austrian Upper Secondary Level Oral Leaving Examination at Colleges for Higher Vocational Education*. Vienna: CEBS (Center for Vocational Languages).

Purves, Alan C. (1971): The Evaluation of Learning in Literature. In: Bloom, Benjamin S.; Hastings, J. Thomas & Madaus, George F. (eds.): *Handbook of Formative and Summative Evaluation of Student Learning*. New York: McGraw-Hill, 699-766.

Richards, Jack C. (2013): Curriculum Approaches in Language Teaching: Forward, central and backwards design. *RELC Journal* 44: 1, 5-33.

Schneider, Günther & North, Brian (2000): *Fremdsprachen können: was heisst das? Skalen zur Beschreibung, Beurteilung und Selbsteinschätzung der fremdsprachlichen Kommunikationsfähigkeit. Nationales Forschungsprogramm 33*. Chur, Zürich: Rüegger.

Stathopoulou, Maria (2013): The Linguistic Characteristics of KPG Written Mediation Tasks Across Levels. In: Lavidas, Nikolao; Alexiou, Thomaï & Sougari Areti, Maria (eds.): *Major Trends in Theoretical and Applied Linguistics: Selected Papers from the 20th ISTAL*. London: Versita de Gruyter, 349-366.

Swain, Merrill (2006): Languaging, Agency and Collaboration in Advanced Language Proficiency. In: Byrnes, Heidi (ed.), *Advanced language learning: The contribution of Halliday and and Vygotsky*. London, New York: Continuum, 95-108.

van den Akker, Jan; Gravemeijer, Koeno; McKenney, Susan & Nieveen, Nienke (2006): *Educational Design Research*. Routledge: London.

van Lier, Leo (2004): *The Ecology and Semiotics of Language Learning*. Dordrecht: Kluwer Academic.

van Lier, Leo (2010): The Ecology of Language Learning: Practice to theory, theory to practice. *Procedia Social and Behavioral Sciences* 3, 2-6.

Wenger, Etienne (2006): *Communities of Practice: A brief introduction* [Online: http://wenger-trayner.com/introduction-to-communities-of-practice/, 18.08.2018]

Zarate, Geneviève (2003): Identities and Plurilingualism: Preconditions for the recognition of intercultural competences. In: Byram, Michael (ed.): *Intercultural Competence*. Strasbourg: Council of Europe, 84-117.